# EPIC2020

Proceedings of a global, virtual conference powered from Melbourne, Australia

**19–24 October 2020**



SCALE

EPIC

**EPIC promotes the use of ethnographic principles to create value for people and organizations.**

EPIC people bring conceptual leadership and practical skills to the problems and opportunities faced by organizations, markets, and society. We are practitioners who work to ensure that innovation, strategies, processes, and products are anchored in deep understandings of people and their everyday lives.

EPIC is a nonprofit driven by volunteers and supported by members and sponsors. We draw on the immense expertise of our global community to learn, connect, and develop as professionals, amplifying our impact together. We collaborate year-round on epicpeople.org and at our annual conference to share expertise and resources, experiment, push each other to improve, and make change happen.

Our conference program is developed by independent committees. The Catalysts, Case Studies, and PechaKucha published in these proceedings are selected through competitive, anonymous peer review.

Join us! epicpeople.org/membership

**epicpeople.org | info@epicpeople.org | @epicpeople_org**

## EPIC BOARD

Nichole Carelock
*Tech Talent Project*

Martha Cotton,
Board Treasurer
*Fjord*

Gary Gebhardt
*HEC Montréal*

Simon Roberts,
Board President
*Stripe Partners*

Carol Zatorre
*Kyvo*

## EPIC STAFF

Jennifer Collier
Jennings, Director of
Content & Comms

Rita Denny
Executive Director

Eliza Parish
Membership &
Operations Manager

## SPONSORS
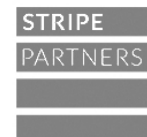
**PLATINUM**

Google        ATLASSIAN

**GOLD**

intel        Spotify

**SILVER**

convo strategic conversations        GEMIC REIMAGINING MARKETS        Multilingual Connections

R&D*        salesforce Research & Insights        STRIPE PARTNERS

# Conference Organization

**CONFERENCE CO-CHAIRS**

Tania Lewis
RMIT University

Alexandra Zafiroglu
Australia National University

**CATALYSTS**

*Erin Taylor, chair, Canela Consulting
Laura Cesafsky, Lyft
Afra Chen, Ipsos
Charlie Cochrane, Jump the Fence
Tom Lee, UC Berkeley
Guillaume Montagu, _unknowns
Heli Rantavuo, Spotify
Josef Wieland, Gemic

**CASE STUDIES**

*Sophie Goodman, co-chair, Deloitte
*Bec Purser, co-chair
Lindsay Ferris, OKCupid
Tom Hoy, Stripe Partners
Ramon Lobato, RMIT
Scott Matter
Chad Maxwell, Kelly Scott Madison
Paul Merrell, Meld Studios
Elena O'Curry
Ovetta Sampson, Microsoft

**GLOBAL COHORT CHAIR**

Nichole Carelock, Tech Talent Project

**PECHAKUCHA**

*Kat Ekberg, co-chair, Westpac
*Evan Hanover, co-chair, Conifer
Nick Agafonoff, Real Ethnography
Frank Romagosa, HiRoad/BlueOwl
Tabitha Steager, Workday
Carrie Yury, Fjord

**REGIONAL CHAIRS**

Charlie Cochrane, Jump the Fence
Marta Cuciurean-Zapan, IDEO
Tom Rowley, Stripe Partners

**GRADUATE COLLOQUIUM**

Jay Hasbrouck, Facebook
Annette Markham, RMIT

**NETWORKING**

Matthew Bernius, Code for America
Tahni Candelaria, Independent

**PROCEEDINGS & ACCESSIBILITY**

Gregory Weinstein, CVS Health

## STUDENT VOLUNTEERS

Andrew Glover, RMIT University
Nicholas Hill, RMIT University
Teffera Girma Teffera, Australia National University
Charlotte Bradley, Australia National University
Memunat Ibraham, Australia National University
Rachel Aalders, Australia National University
Gitika Saksena, SOAS University of London
Abhishek Mohanty, SOAS University of London
Clarice Princess Mudzengi, University of Zimbabwe
Kim Munro, RMIT University
Shriram Venkatraman, IIIT-Delhi
Belinda Glove, RMIT University
Viviane Clement, Texas A&M University
Alen Agaronov, Harvard University
Meg Drechsler, Adelaide University
Philana Woo, University of Chicago
Kelly Chan, RMIT University
Harry Snow, University of New Mexico
Jia-Hui, Lee MIT
Megan Kelleher, RMIT University
Minsoo Bae, Yale
Varshini Balaji, Indiana University

# Contents

## Humanizing Scale

## Scaling Ethnography

## Transcending Time and Space

**Post-Human Scale**

**Scaling Research**

**Empathy Writ Small & Large**

**Dissent at Scale**

**Impact at Human Scale**

**Isolation and Connection**

**Overcoming**

**THEMATIC SESSION**

# Challenging Bigger, Faster, Easier, Broader

In this session, we consider how ethnographic perspectives illuminate the importance of goals beside or alongside 'bigger, faster, easier, broader' and how we re-orient stakeholders to new goals while still creating business value. We'll discuss shared challenges and strategies for building corporate models of scale that can speak to and learn from alternative goals and metrics.

Session Curators: Afra Chen, Lisa Kleinman, Carry Yury

PECHAKUCHA

# Repurposing Risograph Machines

## The Allure of Small-Scale Printmaking in the Era of Big Data

JOYCE S. LEE, *University of California, Berkeley*

*Paper documents are increasingly being replaced with digital files, infinitely replicable for seemingly no cost. Yet I've always felt the pull of paper, with a personal affinity for physical books and a background in magazine production. Through my recreational publishing practice, I learn of the "riso" or risograph, a duplicating machine increasingly adopted by Bay Area artists and technology corporations alike. Upon first glance, most risograph models resemble familiar Xerox machines, with their boxy, gray exteriors, protruding buttons, and hinged tops that cover glass beds for scanning.*

*Through my own experiences and interviews with subject matter experts, however, I come to understand the allure of the risograph: its temperamental nature as an analog machine and the uniquely "human" quality of the prints it renders. I posit the risograph's popularity is a response to technological advances and resulting societal changes, acting as a reprieve to digital modes of aesthetic and community engagement. The rise of the risograph machine thus suggests both the appeal and the limits of scale: its origins as a low-cost printer among artists highlight the draw of multiplicity, whereas its adoption among tech companies, ironically, suggest the limits of seemingly infinite content and growth.*
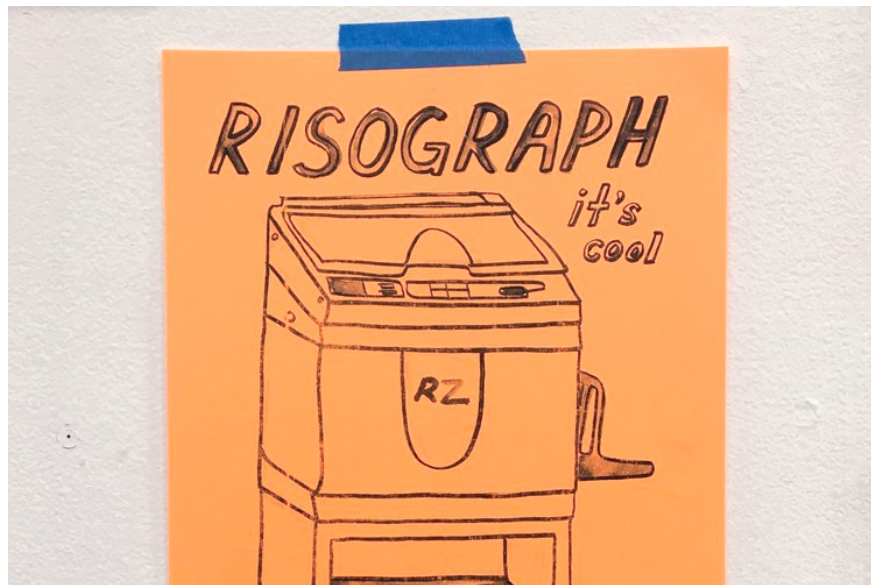
"RISOGRAPH: it's cool" © Amy Burek (Courtesy of Amy Burek)

**Joyce S. Lee** (joyceslee.com) is a user experience researcher based in Oakland, California. In her free time, she publishes and distributes zines; her work has appeared at events around the world including the New York Art Book Fair, the CtrlZ.AI zine fair in Barcelona, and many more festivals across the west coast.

# Who Gets to Define Success?

## Listening to Stories of How People Value Firefox to Redefine Metrics and Revive a Decommissioned Product

GEMMA PETRIE, *Mozilla Firefox*
JENNIFER DAVIDSON, *Mozilla Firefox*

*Challenging measures of scale is possible through listening to stories of how people value a product, and envisioning ways to measure success beyond typical metrics like Monthly Active Use (MAU) or Daily Active Use (DAU).*

*Understanding what people value is somewhat complex for a product like Firefox because people might use Firefox every day without thinking much about it. In this case study, we detail how we used Futures Thinking and participatory design methods to elicit stories of how people value Firefox.*

*This case study demonstrates that a relatively small number of meaningful ethnographic insights can be powerful enough to influence business strategy. By creating the space for listening to stories and encouraging stakeholder involvement, we were able to make the case to save one of our mobile browsers, Firefox Focus, despite its lack of scale.*

## CONTEXT

Over 200 million people use the Firefox web browser every month (Mozilla 2020a). While this works out to less than 10% market share (Statcounter 2020), Firefox has arguably achieved classic definitions of scale. However, the number of people who use Firefox each month has been decreasing over time (Mozilla 2020a) and many efforts at Mozilla, the company behind the Firefox browser, have sought to understand and stop this decline. Measuring success through how a product scales is commonplace. Scale is often assessed through things like Monthly Active Use (MAU) and Daily Active Use (DAU), and entire communities exist to simply increase the growth curve of those numbers (GrowthHackers 2020). We respect the need to measure Mozilla's impact and scale through the sheer number of people who use the Firefox browser, yet as ethnographers, we also know that the reasons behind product choice and usage are often more complex than numbers alone can illustrate.

This case study will discuss a research effort aimed at getting to the heart of a fundamental question: How do people describe the value they get out of Firefox? We hypothesized that by better understanding how people describe the value they get out of Firefox, we would be able to better inform how to measure our success as a company and encourage our leaders to complement traditional measures of scale with more human-centered metrics. This question may strike readers as almost too fundamental. After all, shouldn't product value be well understood after being in the market for over 20 years? But

commonplace products like a web browser present unique challenges for ethnographers. The role of a web browser is almost akin to a utility–it is deeply domesticated (Haddon 2005) into people's lives. People may use Firefox every day without thinking much about it.

Another unique challenge for Mozilla is that the usage data to understand how people use Firefox is often nonexistent. Mozilla practices very limited data collection Our data practices are aligned with our mission[1] and we do not collect information about the content people visit on the web, or spend our resources building usage profiles to sell to advertisers (Mozilla 2020b, Mozilla 2020c, Mozilla 2020d). Often, user research is the only opportunity our organization has to understand the content people seek out and their workflows within the browser. For these reasons, we knew we needed to ground our research approach in methods that would help us dig deep and really get at the root of how people value Firefox.

The genesis of this project came out of two related, but distinct efforts. The first effort was led by our Data Science team and sought to review our current in-product metrics in order to better understand how to interpret our usage numbers and expose any gaps. Our User Research team consulted on that project and followed along with the results. That project exposed a gap in our metrics understanding, where there was limited qualitative explanation of usage numbers that were grounded in ethnographic research.

The second effort, led by a cross-functional research team, aimed to gain a top-down view of value by asking our senior leaders how they would define the value of our products. Perhaps unsurprisingly, we found that not every leader had the same answer. There was a lack of alignment around who our products' primary audience is and how we address people's needs in our products.

The authors were each involved in one of these projects, and as often happens with foundational, ethnographic work, we proposed a study that was not previously on our roadmap by identifying an opportunity to align these efforts and explore the gaps we were observing. We hypothesized that better understanding the nuanced ways that people talk about the value they derive from using our products could help us define new, human-centered metrics to measure our success and scale against. We knew it was time to get an "outside in" perspective to help better inform our internal narrative, and ultimately help our organization make better product decisions.

## METHOD

To overcome the challenges around investigating value in a domesticated, routine product we knew we would need to develop a mixed method approach that included interactive activities, and not rely solely on something like a retrospective interview.

The research proposal was completed in August 2019, and the research itself was conducted in late September & October 2019 for Desktop and February 2020 for Mobile. The last report out related to the research was in April 2020.

### Research Activities

The research included three phases: a diary study, remote interviews, and an in-person workshop.

The diary study took place over three consecutive days where participants reflected on their use of Firefox each day. We aimed to get a foundational understanding of how these

particular participants use Firefox. Additionally, we wanted to get participants in the mindset of actually thinking about Firefox and how they value it. To get people in this mindset of thinking about Firefox, participants responded to a pessimistic scenario, inspired by Futures Thinking (Textor 1980), asking how their day would be impacted (or not) if Firefox wasn't available that day.

We used the remote interviews to begin to build rapport with individual participants before the workshop and to learn about participants' history with Firefox. Responses from the diary study were also clarified during the remote interviews.

The third, and main part, of the study was a two and a half hour in-person workshop (Table 1). Each workshop involved five to six participants, and two to four Mozilla employees. The workshop relied on both the remote interviews and diary study to gain a basic understanding of participants and their use of Firefox before diving deeper.

Table 1. Workshop Agenda: an interactive workshop that included a range of activities.

| Length (minutes) | Activity |
| --- | --- |
| 10 | **Settle in.** Get name tag, write pronoun on it. |
| 5 | **Quick introductions.** Why we're here today, introducing workshop leads. |
| 25 | **Warm up.** Longer introductions. |
| 20 | **Optimistic scenario building.** |
| 20 | **Pessimistic scenario building.** |
| 10 | **Break.** |
| 30 | **Metric scenario building.** |
| 10 | **Value prop evaluation.** |
| 15 | **How Firefox should measure success.** Pitch videos. |
| 5 | **Wrap up.** |

The workshop was grounded in Futures Thinking (Textor 1980), where we elicited optimistic, pessimistic, and "normal" scenarios from participants (Figure 1). As mentioned earlier, someone could use Firefox without really thinking about it. So, Futures Thinking was particularly appropriate in this case, to enable participants to think deeply about how Firefox is valuable to them. The workshop resulted in many, varied real and imagined scenarios that included not only how they value Firefox, but also how they feel about that value.

Figure 1. Optimistic Scenario Building Worksheet. Positive retrospective worksheet to elicit conversations about how people value Firefox.

Given the gaps we observed while working with our Data Science team around a lack of qualitative understanding of our usage metrics, we used these workshop exercises to reflect on the stories behind Firefox usage. For example, we asked participants to describe a time when they searched more in Firefox, to qualitatively understand a metric of "amount of searches" (Figure 2).

Who Gets to Define Success? — Petrie & Davidson

Figure 2. Amount of Searches Worksheet. Retrospective worksheet to elicit conversation about what stories might be behind our metric of "amount of searches".

At the end of the workshop, inspired by participatory design methods (Stephen 2012), we "showed our cards", and asked participants to create a pitch video about how they thought Mozilla leaders should measure the value people get out of Firefox. Each participant was video recorded giving a short (1 minute or less) fictional pitch to Mozilla leadership. The participatory method of the pitch video provided our team with new ideas. For example, a participant in Berlin focused on a theme of security, and whether or not people understand security, privacy, and what data is collected from Firefox:

> *"I would base my assumptions on usage figures, i.e. usage period, age. I'd also argue that Mozilla is limited to the most important features and remains minimalistic and clear – especially on the phone. It would also be important that security is well understood by users – especially younger folks. It should be easy to understand security settings and stay informed about which data is collected."* - Participant in Berlin

While our organization tracks security bugs, and reduces errors as much as possible with each Firefox release, we do not use security and its understanding as a top-level success metric for our products. Measures related to privacy and security would be wholly aligned with our mission[1], and we are grateful to the participant for their ideas.

## Location

Our research took place in six locations: Berlin, Chicago, Portland (Oregon), Seattle, Taipei, and Vancouver (Canada). The locations were decided using a variety of factors, including 1) where Firefox market share is, 2) where stakeholders are located, 3) where researchers are located, and 4) where we have not done ethnographic research in the recent past. For 2), note that stakeholders are distributed around the globe as Mozilla has many remote workers.

All research materials and activities were conducted in English for locations in North America, in German for Berlin, and in Chinese for Taipei. In Berlin, there was a simultaneous interpreter present for the remote interviews and workshops to translate to and from English and German, as many people in Berlin communicate in English. Research notes, diary entries, and pitch videos were translated back into English for analysis.

## Participants

In total, there were 61 participants across all the locations. The participants remained the same during each of three phases, to help us get a deeper understanding of how they value Firefox. To explain, we ran two workshops in Vancouver, and there were five to six participants in each workshop. Those same five to six participants also took part in the diary study and remote interview.

Participants were recruited using a professional recruiting agency and through an in-product invitation. In-product recruiting for in-person research was relatively novel at Mozilla, so only a handful of participants were recruited that way, to test the capability. It's outside of the scope of this case study, so suffice to say, in-product recruiting worked.

All participants were required to have and use Firefox on either a desktop computer, a mobile device, or both. Participants were all 19 years or older, and spoke either conversational English, German, or Chinese, depending on where they were located. As Firefox is used by a range of individuals, we aimed to get a diverse representation of people, on the following axes: self-reported weekly hours of using Firefox, operating systems, job status, industry (if applicable), gender, age, early adopter status, educational background, income, race/ethnicity/tribe. Race/ethnicity/tribe were only asked in the US and Canada due to regulations.

These methods are explained for other researchers to have insight into how they might replicate an interactive discussion with people who use their products to learn how they value those products, and how they would suggest the business measure success.

## INSIGHTS ABOUT OUR PROCESS

### Stakeholder Involvement

Our six workshops were run globally. We leveraged facilitation skills from designers and researchers across Mozilla, even outside of our team. This was a large investment to ask of our organization, so we knew successfully launching this study would require buy-in from a wide variety of stakeholders. Because of this, we spent more time in the research proposal phase than usual (over a month), soliciting feedback and incorporating changes into our

proposal. Our recommendation to other ethnographers is: Do not underestimate the proposal phase. Even though we had the budget on our team to do the work, we knew we needed people to not only see this work, but see it as a priority and something worthy of their attention. The proposal included what the organization could gain from the work. For example, we argued that this work could provide empirical grounding for current and future metrics.

Mozilla stakeholders appreciate and request research, oftentimes more than what our small team can handle. Barriers for stakeholder participation were not related to valuing research in general, but rather more practical barriers like prioritizing participating in this research project compared to other day-to-day work. So, we worked diligently to get stakeholders involved in a hands-on way with the research activities. Instead of sending out a general call for note-takers and observers, we reached out to specific individuals that we wanted to encourage to participate. We explained why we thought they might be interested, and why we wanted them to experience this work first-hand. In some cases, we even chose locations that were close to our most senior stakeholders in order to increase the likelihood of their participation, since we knew that limiting long distance travel would enable more of them to join our field team.

Additionally, we set up a half-day after each workshop to dig into analysis with stakeholders. We set up detailed spreadsheets ahead of time to make coordinated analysis possible with multiple locations conducting this work simultaneously (Figure 3). The spreadsheets allowed for multiple field teams to enter data at the same time in a structured way, which greatly simplified our more formal analysis work later on. Finally, after we created a draft report, we piloted our talk with just our stakeholders to get their thoughts and feedback before sharing it with a larger audience. We've found that giving stakeholders a preview often means that they feel more confident contributing to the discussion during larger share outs.



Figure 3. Group Analysis Spreadsheet. A section of our group analysis spreadsheet showing how we managed data from multiple locations simultaneously.

Who were our stakeholders? We focused our stakeholder recruitment mostly on Program Managers and Program Directors. We also had strong support from Data Science, who helped us thoughtfully consider the potential impact of the measurements our participants proposed (i.e. Could we measure what was suggested? How would we measure it?). We were excited to include a few new hires as field team members (some in Data Science, some in Program Management), since we believe that the experience of a user research study, and hearing directly from people using our products, is an excellent way to onboard new colleagues.

## INSIGHTS FROM PARTICIPANTS

This research helped us better understand the value of our products by focusing on stories behind people's needs and the workflows people use to accomplish their goals. We learned that the top-of-mind most valuable activities that participants use Firefox for are:

**Table 2. Valuable Use Cases in Firefox Desktop and Mobile.**

| Firefox Desktop | Firefox Mobile |
|---|---|
| • Performing information seeking activities.<br>• Engaging in various forms of entertainment.<br>• Communicating with friends, family, and colleagues.<br>• Accomplishing personal and work-related tasks. | • Performing quick informational tasks.<br>• Engaging in various forms of entertainment.<br>• Holding information for later. |

*Participants shared their most valuable ways they use Firefox, and they fell into the categories above.*

Using our Futures Thinking exercises, where we elicited optimistic, pessimistic, and "normal" scenarios from participants (Figure 1), we learned that when Firefox works well, participants feel "productive", "happy", "efficient", "in the flow", and "normal." Imagining when Firefox doesn't work well (like losing all their saved history, passwords, and bookmarks), participants felt "indifferent, "angry", and "stressed."

We also learned how participants felt in scenarios that were related directly to our metrics (Figure 2). It showed us something we, as ethnographers, often feel–that positive "hockey-stick" style growth or usage, is not always a "good thing" for someone using a product. For example, one participant in Vancouver, Canada described preparing for a race in Whistler, Canada called the Donut Dash. He described researching the race rules, FAQs, the registration fee, the registration form, and the race course. All of these details were open in different tabs in his Firefox browser. He started his search excited to sign up for the race, but ended up feeling overwhelmed by the number of tabs he had open and how hard it was to keep track of all the different information he needed to understand. This participant

described this experience of "information overload" as something that ultimately made him feel uninformed, unprepared, and unmotivated. The number of tabs correlates with a metric around intensity of use, which as a business is viewed positively and is valuable for revenue. However, for this participant, his story shows how having so many tabs open was overwhelming to him.

As we mentioned earlier, one of our goals was to introduce new human-centered metrics to how our organization measures success. Input from our participants resulted in specific recommendations including: A search satisfaction metric, and a metric to measure if people using Firefox are in the psychological state of "flow" (Wikipedia 2020).

## RESEARCH IMPACT

### Use Case Articulation

Earlier, we mentioned that this study was not just about what people do with Firefox, but how they value it. However, we were initially surprised that one of the stickiest results from this study was, in fact, what people do with Firefox. Upon reflection we understood that, because of our mission-driven commitment to limited data collection, our engineering-led organization tends to focus more on the mechanics of maintaining an open-source browser, rather than investigating what people might be using the browser for. Ultimately, it makes sense that our organization would latch on to this high-level overview of how people are using Firefox on desktop and mobile devices in the absence of comprehensive personal data collection.

While this result told us what participants used Firefox for, it also showed us what kinds of internet activities were most **important to them** (Table 2). This data came from an introductory activity during the workshop where we asked people to write or draw at least one important thing they do on the internet. We followed this exercise with a discussion where participants explained their choices and built on what other participants were sharing. The researchers then coded all of these examples against a primary list of internet workflows that our team has developed over our years of research to categorize them into high-level themes.

This use case articulation ended up inspiring a variety of mixed method efforts in our organization. Other members of our research team, primarily Rosanne Scholl, used this information to design multiple surveys fielded across thousands of individuals to see how these categories "rate" at scale. The results of these surveys showed that these categories were highly durable and effectively described the high-level activities that people who use Firefox engage in with the browser. One interesting survey finding was that "education" emerged as a frequently cited additional important use case for desktop in our surveys, which coincided with the beginning of the COVID-19 pandemic and a massive cultural shift toward online learning. The results from this case study also inspired a design sprint on the topic of entertainment, a topic that Mozilla had not previously dedicated many resources to.

**Saving Firefox Focus**

This research effort was a large investment for our organization, but one that has been widely regarded as an important piece of foundational research for Firefox. We presented a large number of recommendations to various teams, and as often happens, the findings that ended up gaining traction were not necessarily tied to our original intent, but are no less important. We went into this research hoping to inspire our organization to describe and instrument new metrics to measure our products and their success. In fact, Data Science's involvement in our research strengthened our recommendations related to human-centered metrics. But, in addition to the unexpected impact of our participant use case articulation, the biggest success coming out of this work is that we were able to save a decommissioned product: Firefox Focus.

Firefox Focus is a specialty mobile browser designed around privacy and simplicity. Focus automatically blocks a wide range of online trackers and makes it easy to erase history, passwords, and cookies with a single button, ensuring people won't be followed by things like unwanted ads (Figure 4). Focus has a relatively small number of people who use it and does not have a measurable impact on Mozilla's revenue. As a result, a decision was made in early 2019 to sunset Firefox Focus due to resource constraints in an effort to simplify our product portfolio. The sunset decision was reversed because of our research.
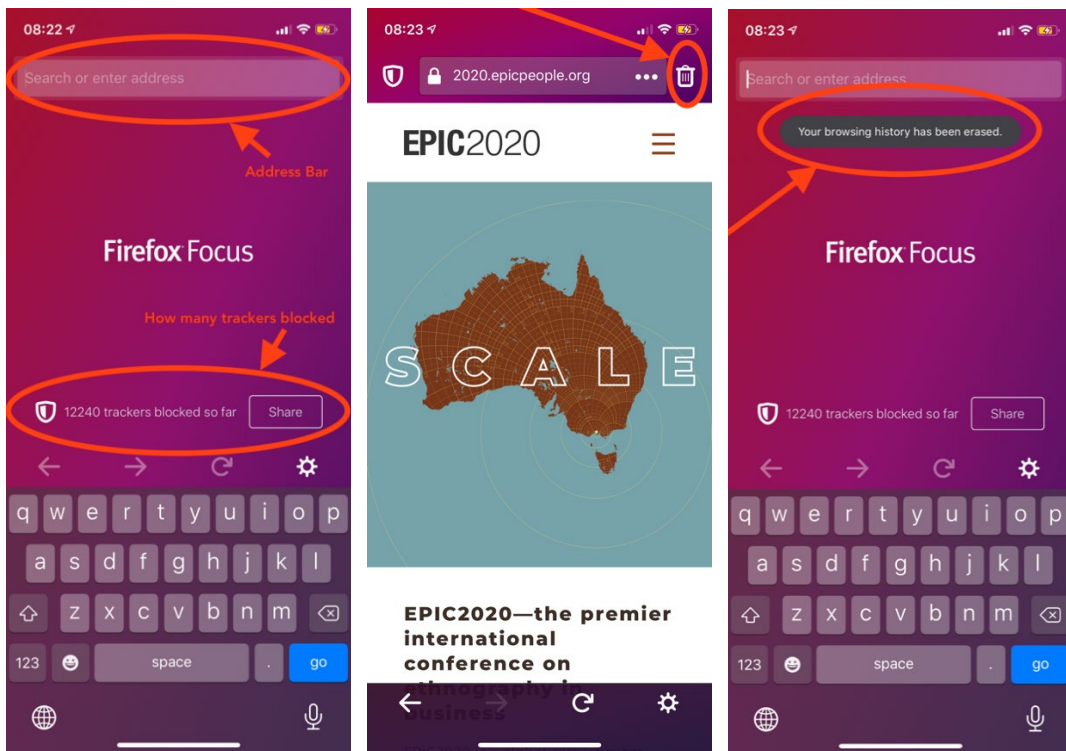


Figure 4. Firefox Focus. From left to right: 1) Firefox Focus Home Screen that shows an Address Bar and how many trackers have been blocked; 2) Firefox Focus when visiting a webpage, including the "trash" icon to the top right; 3) After selecting the "trash" icon, Firefox Focus shows that browsing history has been erased.

We were able to show through this study that despite its relatively small usage base, Focus is often used alongside another Firefox mobile browser and plays a critical role in some people's workflows for specific use cases. For example, a preschool teacher in one of our Seattle workshops described themselves as a long-term user of Firefox. They use Firefox on both their desktop and their Android device, and they also said they use Focus for specific tasks. They do a lot of research on their mobile device related to their hobbies – things like gardening and vegan cooking. This participant is also quite politically active and they described how they switch to Focus for their political research because they, "Don't always want things recorded" (Figure 5). We heard similar things from other participants who used Focus in our workshops. Focus was often present alongside another mobile browser and used for specific kinds of tasks–sometimes for content that was sensitive in nature, but other times for quick one-off searches because participants liked starting each session fresh and knew there was some information they didn't need to retain.
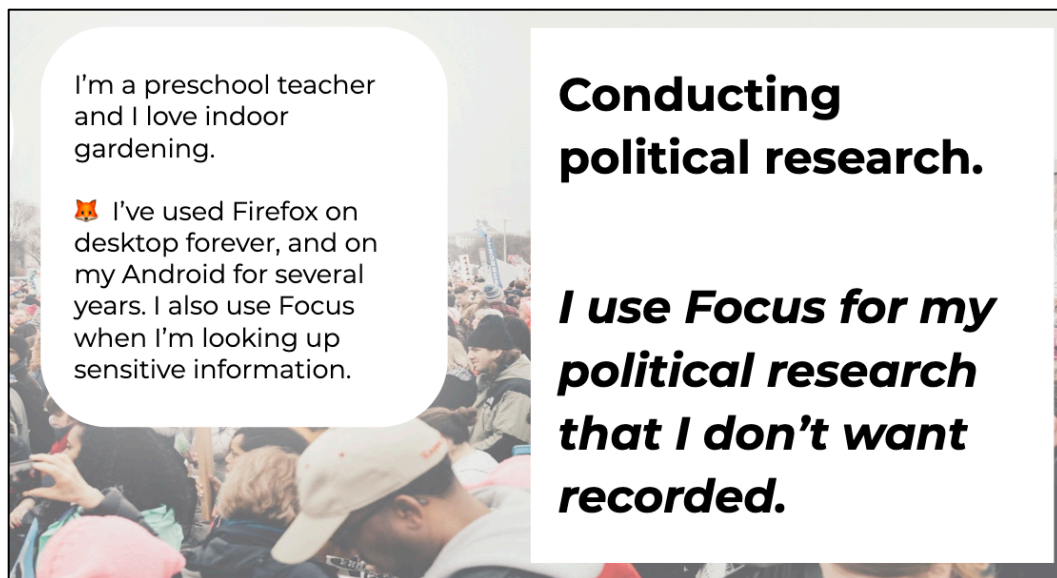


Figure 5. Using Firefox Focus. Excerpt from internal presentation to stakeholders.

After multiple report-outs of this research work, and digging through past Firefox Focus research primarily conducted by team member Alice Rhee, our mobile business strategy was changed to not only use the calculable metrics of daily active use or number of downloads, but to also include a deep consideration for the people who already use and love Focus. We attribute the change in decision in part to the fact that we had the support of two senior PMs in this product space who were part of our field team. These individuals were able to hear these stories from participants first-hand and debrief with our field team after our workshops. We were able to make the case to save this unique product in spite of its lack of scale, and ultimately alter our organization's view on how we can define the success of our products. Our exact recommendation for Focus was:

*"**Bring focus back to Focus:** Continue to support Focus. Participants that were using Focus were often using it alongside another browser for specific tasks and valued the simplicity of the experience. Can we get folks who use our Firefox mobile browser to **also** use Focus for their quick searches?"*

The above recommendation did not refer to usage metrics, like "how many people use Firefox Focus compared to our flagship browser." Instead, we explained how people value Firefox Focus. We believe this focus on value helped us influence the decision to keep Firefox Focus in our product suite.

After giving presentations all over the organization, including a lightning talk that inspired colleagues to create a particularly fun Zoom online meeting background (Figure 6), a decision was made to keep Focus around.



Figure 6. Save Firefox Focus. Zoom background used internally to promote strategy change.

## CONCLUSION

This case study explains how our original intent was to re-define, or add to, our organization's current ways of measuring success. However, by listening to individuals' stories through methods inspired by Futures Thinking and participatory design, and involving our stakeholders during every step from planning to analysis, we had another outcome. Decision makers used the stories they witnessed first-hand about how participants deeply valued Firefox Focus to revive this decommissioned product. We urge other ethnographers to use their research to challenge, question, or complement typical measurements of scale, listen to the people who use your products, ask for their opinion on how to measure success, and as always, bring decision makers into the field with you.

**Gemma Petrie** is a Principal Researcher at Mozilla Firefox. She enjoys conducting research using a wide variety of methods, building and leading teams, and working to increase the

value of research in organizations. Gemma has a BA in Philosophy from Reed College and an MLIS from the University of Illinois Urbana-Champaign.

**Jennifer Davidson** is a Staff Researcher at Okta, previously at Mozilla Firefox. Aside from being a mixed methods researcher in industry, she enjoys teaching at universities, non-profits, and vocational schools. Jennifer has a PhD in Computer Science, with an emphasis in Human-Computer Interaction from Oregon State University.

## NOTES

1. "Our mission is to ensure the Internet is a global public resource, open and accessible to all. An Internet that truly puts people first, where individuals can shape their own experience and are empowered, safe and independent."

## REFERENCES CITED

GrowthHackers. 2020. *GrowthHackers* website. Accessed September 27, 2020. https://growthhackers.com/posts.

Haddon, Leslie. 2006. "The Contribution of Domestication Research to In-Home Computing and Media Consumption." *The Information Society: An International Journal* 22 (4): 195–203. doi: 10.1080/01972240600791325.

Lindsay, Stephen, Daniel Jackson, Guy Schofield, and Patrick Oliver. 2012. "Engaging older people using participatory design." In *Proceedings of the SIGCHI conference on human factors in computing systems*: 1199–1208. doi: 10.1145/2207676.2208570.

Mozilla. 2020a. "Firefox Public Data Report." *Mozilla* website. Accessed September 27, 2020. https://data.firefox.com/dashboard/user-activity.

Mozilla. 2020b. "Lean Data Practices." *Mozilla* website. Accessed September 27, 2020. https://www.mozilla.org/en-US/about/policy/lean-data/.

Mozilla. 2020c. "Mozilla Privacy Policy." *Mozilla* website. Accessed September 27, 2020. https://www.mozilla.org/en-US/privacy/.

Mozilla. 2020d. "We're Building a Better Internet." *Mozilla* website. Accessed September 27, 2020. https://www.mozilla.org/en-US/mission/.

Statcounter. 2020. "Desktop Browser Market Share Worldwide." *Statcounter* website. Accessed September 27, 2020. https://gs.statcounter.com/browser-market-share/desktop/worldwide.

Textor, Robert B. 1980. *A handbook on ethnographic futures research*. Stanford, CA: Stanford University Press.

Wikipedia. 2020. "Flow (psychology)". *Wikipedia* website. Accessed September 27, 2020. https://en.wikipedia.org/wiki/Flow_(psychology).

**CATALYST**

# Everybody's a Winner

## A Study on How Unidimensional Scaling Up as an Entrepreneurial Rite of Passage Is Beginning to Be Resisted in India's Startup Capital

GITIKA SAKSENA, *LagomWorks Consulting & SOAS University of London*
ABHISHEK MOHANTY, *LagomWorks Consulting & SOAS University of London*

*Taking India's startup capital Bangalore as its field, the paper researches the absence of conventional scale as a potentially positive emic experience for the entrepreneur. The study leverages a mixed methods approach, employing semi-structured interviews with select entrepreneurs, employees, investors, advisors, and staff from startup-incubators, participant observation at both startups and startup-incubators, textual analyses of business literature, semiotics of popular culture as well as auto-ethnographic reflection by the authors themselves on account of having co-founded a company in Bangalore in 2018, therefore establishing their positionality as 'an-other' (Sarukkai 1997, 1408), by 'thick participation' (Samudra 2008, 667). The authors examine the current assemblages within the startup ecosystem, to demonstrate that even the dominant and conventional notions of scale have begun to demonstrate multidimensionalities. At the same time, the authors advocate a case for tethering at different points of the scalar continuum as an alternative model of entrepreneurship. The authors share ethnographic evidence from their research on startups that are increasingly beginning to explore potentialities and innovation at the existing point of the scalar continuum through an exercise in consolidation and tethering. Finally, the authors advance the proposition that a quest for value is not necessarily resistant to scale, and concomitant streams of nonscalability as lines of flight existing along the periphery of incumbent structures, always carry the possibility to escape and thus, create potentialities for creativity and disruption.*

*Keywords: Scale, Startups, Entrepreneurs, Interscalarity*

> '*The facilitation under the Startup policy is intended for only technology based startups, i.e. one that creates a technology-based service or product or uses technology for* enhancing functionality or reach of an existing product or service' (Startup India 2020).

> '*[A technology startup is]... an entity working towards innovation, development, deployment, and commercialisation of* new products, processes, or services *driven by technology or intellectual property*' (NASSCOM Zinnov 2018).

## A TALE OF THREE CO-FOUNDERS

On a pleasant Bangalore summer evening in 2019, is when we first met with Saurav, Varun, and Abhay, the young co-founders of an AI-powered startup focusing on helping children from small-town India improve their spoken English skills. Sipping on *masala chai*,

they recounted how it was the near-debilitating insecurity they felt when at university, around their own inabilities to converse in English as fluently as their city-educated classmates could, that spurred them to found their startup, which we will call *Chalk Test*.

'Our *shared, personal experiences,*' Saurav explained, gesturing towards his co-founders, also classmates from university, *'have shaped the direction in which the three of us want to move our company. The problem of not being able to speak in English and thus suffer a significant erosion of confidence, is severe. Importantly, it is not so much a function of the family's income, as it is about the resources and opportunities which children from smaller towns and villages lack.'*

We talked about Bangalore's thriving startup environment. The city continued to be celebrated as the 'top tech startup hub' in the country, and for being amongst the 'top 3 cities globally for …[the]… launch of tech startups' (NASSCOM Zinnov 2018, 12). The modest offices we were in was part of a leafy neighbourhood emblematic of the city's reputation, throbbing as it was with ambition, evident from the several many, unmissable startup address plaques and signages, busy dive bars, hip gastropubs, and bustling hole-in-the-wall restaurants.

'There *is a lot of noise in the startup community,*' Saurav continued. *'Everybody talks about funding, the next million, and being featured in the media. All of this makes you question your core beliefs and assumptions. We need to be careful not to be influenced by the noise.'* Vaibhav joined in, *'Social purpose must be balanced with commercial sustainability and scale. We have to manage two contrasting aspects: creating impact and making money. We have not been able to find the balance yet. We also have to pay our staff and team of freelancers. That is why we have plan B, to focus on non-core revenue which can cover expenses and decrease burn. We need specific skills to scale. Plan B helps us buy them. That said, we have to think about how we can monetize the app and breakeven soon.'* This served as a cue for Abhay, the relatively quieter one of the three co-founders. *'We have hundreds of positive testimonials from children who have used the app,'* he tabled. *'This is real impact, and this is what keeps us motivated. After all, an increase in a student's confidence is success for us. Increase in speaking time and frequency are metrics we have begun to track. They show if the student has begun to speak her mind.'* *'We see no difference between a social and a for-profit enterprise,'* quipped Varun. *'Adding social as a term to an enterprise essentially gives us a framework for decision-making. The fundamentals of business are the same. What I am saying is that we need to work on our metrics, create impact, and capture the wider market.'* *'We have to find a balance between cause and commerce in our metrics and measures as well,'* summarized Saurav. *'We do not want to be romantic social entrepreneurs. For while we know that education is a slow and hard business, we fear that we may just be running out of time.'* (Field Notes Extract 2019)

## INTRODUCTION

Our research has its roots in a project commissioned by one of our clients, a Bangalore-based innovations incubator which we will call *InnoCubator*, that sought to explore how certain early-stage ventures in its portfolio could scale-up. Yet more importantly, even if implicitly, it wanted to understand what scaling-up really meant, outside of the unidimensional, *le manuel scolaire* perspective of growth as defined in terms of financial, market and customer metrics. To us, the project-ask was, in and of itself, a notable point of departure. A startup incubator going beyond traditional metrics and seeking an alternative grasp of scaling-up was uncommon, to say the least. For as Tsing (2019, 143) hauntingly advances, while alluding to scale as an exercise in precision, 'there is something disturbingly beautiful about precision, even when we know it fails us'.

Tsing warns us of the dangers of a relentless pursuit of scale-making where 'bigger was always better' as one anchored in expansionism of the kinds which ignores 'meaningful diversity' (2019, 145-146). Our ethnographic encounters with entrepreneurs in Bangalore over the summer of 2019 demonstrate a remarkable grasp on their part of Tsing's cautionary note. As in the case of the co-founders of *Chalk Test*, our conversations were invariably peppered with references to scale and its concomitant notions of scaling up and scalability, almost in the sense of a Durkheimian social fact. Yet just as our three young protagonists simultaneously acknowledged and sought to dialectically negotiate such hegemonic and unidimensional narratives through arguments anchored in alternative 'plan B revenue streams' which allowed them to focus on 'impact', and articulations of the need to balance 'social purpose' with 'commerce' (Field Notes Extract 2019), so did the other entrepreneurs, academics, and industry experts we engaged with.

As our study eventually showed, a small but growing breed of entrepreneurs, investors, and incubators in India was beginning to view failure to scale in the conventional sense, as advancing a Deleuzian glance at an entrepreneurial becoming, and an opportunity to pivot to alternative goals and means of engendering value. Scale was being at once resisted and negotiated to incorporate the non-scalable through inter-scalar articulations and balance, as in the case of our three *Chalk Test* co-founders, in the sense of assemblages comprising 'plan B' and 'non-core' activities on one hand, and 'real impact', 'specific skills' and 'getting the technology right' on the other hand. Or as a seasoned academic we spoke with, pointed out:

> There is no definition of what constitutes scale. Each entrepreneur should decide the framework and time period to achieve scale based on his or her priorities. Everyone need not become a unicorn or float an IPO. Keeping your head above water for a long period of time may also be sufficient for someone. (Field Notes Extract 2019)

Carr and Lempert (2016, 8-9) tell us that a meaningful ethnographic approach to understanding scale situates it as 'a practice and process before it is … [a] … product'. Taking India's startup capital Bangalore as its field, this paper researches the absence of conventional scale as a positive emic experience for the entrepreneur, undergirded by the affective imaginaries of passion, independence, and perseverance. The study leverages a mixed methods approach, employing semi-structured interviews with select entrepreneurs, employees, investors, advisors, and staff from startup incubators, participant observation at both startups and startup incubators, textual analyses of business literature, as well as auto-ethnographic reflection by us on account of having co-founded a company in Bangalore in 2018, therefore establishing our positionality as 'an-other' (Sarukkai 1997, 1408), by 'thick participation' (Samudra 2008, 667). In doing so, the study posits that the problematic playing out in India's entrepreneurial zeitgeist is not necessarily a summary rejection of notions of economic or financial value, but a nuanced adoption of balance incorporating recognised key performance indicators (KPIs) of 'bigger, faster, easier, broader' (EPIC 2020), alongside alternative metrics, goals, and measures of value.

In order to develop a deeper understanding of interscalarity, we drew analogous inspiration from Susan Philips' ethnography of how 'legal activity is interscaled in [Tongan] higher and lower trial courts' and in doing so, naturalizes the institutions and their ideologies (2016, 112). She asserts that 'scaling is, after all, a cultural and a semiotic phenomenon' and is characterized by an interdependence of elements, whose intertwining is taken for granted

(2016, 113). Philips outlines that the higher and lower courts are intentionally maintained at distinct levels of scale to ensure relativity or 'the scalar antinomy of "higher" and "lower"' on one hand, and conceptual coherence on the other (2016, 115). This multidimensionality of scale, here expressed as seriousness of the case, plays out along elements such as space (bigger courtrooms), time (longer trials), actors (senior judges), enforcement (of procedures and evidence) and media (both Tongan and English languages being used in the higher courts). The conceptualization and enactment of these hierarchies of scale had its roots in imperialist undertakings, exported by European colonialists as a way of managing complexity and conflict. At the same time, there was a *becoming* of interscalarity as it continued to get influenced by local Tongan circumstances. Philips identifies five interdependent dimensions which distinguish as well as integrate the higher and lower courts to reinforce the scale of seriousness. These dimensions (2016, 121) encompass cultural phenomena including social identities of the key actors (judges and magistrates, plaintiffs and lawyers), privileging of procedure and documentation, use of language (geopolitical and translocal influences of English on production of activity), length of time spent (on evidences and amount of talk), and space of jurisdiction (area of authority, geographical locations, demarcated physical space, and language choices again). In this manner, Philips advances a compelling argument for analysing scale as a social construct, and understanding the cultural aspects that are immanent in its naturalization. Another key insight which can be drawn from the study is that both higher and lower levels along a continuum need to be nurtured to maintain the function of scale. Situating our research in light of Philips' study informed our areas of enquiry. Thus, the social phenomenon of scale as conceived and institutionalized in the startup ecosystem in Bangalore entailed an analysis of multidimensionality. Carrying this notion forward, we defined dimensions or assemblage-constituents as primary actors (entrepreneurs, investors, academia, incubators, mentors), procedures (policy documents, funding eligibility guidelines, incubation competitions), language (media discourses), time (invested in evidence building and success narratives) and space (investments in physical capacity, geographic spread of hubs and startups). This specifically elucidated the scope for review of public culture, as we studied varied texts including government startup policy documents, annual reports on the ecosystem, and published interviews.

Furthermore, analysing these narratives within the paradigm of interscalarity yielded distinct areas of ethnographic enquiry. Our principal research question was to interrogate the emerging dimensionality of scale, as it appropriated and absorbed other institutionalized or existing narratives in its fold. As we have noted earlier, the growth of the startup ecosystem in India did not just yield materialization, in the shape of unicorns, funds and technology, but also sought to achieve the state's priorities of job creation, women empowerment and the development of smaller cities and towns. Second, we asked whether the potentiality of scaling as 'taking wings' or 'achieving escape velocity' presents a risk of *untethering*, in the sense of a weakening of the very foundations of the climate of innovation which the startup ecosystem seeks to thrive in (Startup India 2020). We have observed that scale was increasingly getting entrenched as a qualifier or an entry barrier to participate in the startup ecosystem, be it in terms of the stage of funding (with an increase in late-stage investments), the archetype of the startup founder or entrepreneur, and even the requirement of innovative technology solutions for the realization of social impact. And lastly, we sought to understand if the binary of *tethering* could offer another field of possibility, where

entrepreneurs chose to either entirely opt out of the conventional projects of *high* scale, or continue to negotiate it by concomitantly pursuing *desires*, in the sense of alternate value.

While our research is situated in India, it can be subsequently leveraged for additional comparative studies in other countries, for governments, funders, and organizations in their broader project of supporting entrepreneurship and innovation. Grounded thus in the EPIC community's aim of using 'ethnographic principles to create business value' (EPIC 2020), we intend for the study to realise a contribution to the 'largely missing ... [anthropological] ... research at the level of new ventures' (Briody and Stewart 2019, 142) in the sense of a diacritical mark for ethnographic literature on entrepreneurship.

## SITUATING OUR RESEARCH WITHIN THE POTENTIALITY OF SCALING

*'Breaking $1 billion is a psychological milestone' said Hiten Shah, cofounder of several SaaS companies, including KissMetrics, Crazy Egg, and FYI. 'It indicates that your company is a real force, a business to be taken seriously. It has a cascading effect on the press, investors, and recruitment.' (Sengupta and Narayanan 2019).*

2019 represented a critical epoch for the startup ecosystem in India, with the crossing of significant milestones in the preceding year, and the expectations borne with them. Eight startups had crossed the USD 1 billion in valuation milestone in 2018, and attained the much feted status of a unicorn. The number of new technology startups had seen a year-on-year growth of between 12% to 15%, even as the overall number across the country was expected to surpass 7500 (NASSCOM Zinnov 2018, 3). The 2018 edition of National Association of Software and Services Companies (NASSCOM) report on the 'Indian Startup Ecosystem' placed great emphasis on notions of scale as markers of success, with callouts such as the '*dramatic increase* in number of unicorns, *resurgence* in investments, and *rapid growth* in advanced technology in startup ecosystem' (NASSCOM Zinnov 2018, 3). The report highlighted that the ecosystem was gearing up to attain 'escape velocity'. Projects of global *'expansion'* had been outlined as well, with references to Indian-origin startups registering their presence in markets outside India, as well as 'international startup exchange missions' setting up bases in the country (NASSCOM Zinnov 2018, 3). In this manner, and come 2019, conventional definitions of scale had been institutionalized to characterise the success of the ecosystem, be it through growth in numbers, rise in funding, or an increase in the adoption of advanced technology. An additional manifestation of scale was seen in the projection of the archetypal entrepreneur. The aforementioned NASSCOM report for example, suggested that a *successful* entrepreneur was likely to have a *strong* educational background (as having an engineering, MBA, MS, or PhD degree) as well as prior corporate work experience (of five to ten years) implying *better* networks and skills (NASSCOM Zinnov 2018, 59).

The year 2018 had witnessed another key trend where, while there was an increase in the average funding per deal (by 144%), most of it was directed towards mature startups requiring late-stage investments. This clear preference for scale had led to a year-on-year decline of 18% in funding for seed stage deals. In fact, Debjani Ghosh, the President of NASSCOM, had expressed her concern at the probability that without protection at the seed stage, innovation was bound to get impacted (Variyar 2018). Yet at the same time, a rising heterogeneity in the landscape had also been noted, expressed in terms of the increasing

proportion of women founders, creation of direct and indirect jobs for the economy as well as the emergence of Tier 2 and 3 cities in India as startup hubs (NASSCOM Zinnov 2018, 59).

For our research, we chose Bangalore, the busy capital of the southern Indian state of Karnataka and the country's Silicon Valley, as our field. Officially now known as Bengaluru, the city enjoyed (and continues to enjoy) the position of being not only the primary but also the fastest growing startup hub in India. In 2018, Bangalore was home to one-fourth of the total number of technology-startups in the country. The city was the fulcrum for the Karnataka Startup Policy 2015 - 2020, which sought '*to give wings* to startups in the state through strategic investment & policy interventions by leveraging the robust innovation climate in Bengaluru' (Startup India 2020). Although limited in its scope to technology-led startups alone, the policy's goals were reflective of the varied aims that a scaling-up of the ecosystem could facilitate. By advocating the growth of twenty thousand technology-based startups in Karnataka, the state government's goals were effectively looking at the creation of 1.8 million jobs, galvanizing a startup funding investment of INR 20 billion, and generating at least twenty-five innovative technology solutions in areas of public welfare, such as health, food security, clean environment, and education (Startup India 2020). In this manner the city of Bangalore, as an assemblage of actors, policies, spaces, and media, offered an inimitable opportunity to build ethnographic evidence for developing an epistemology of the implications and perceptions of scale in entrepreneurship.

Subsequently, our ethnographic enquiry leveraged participant observation as well as semi-structured interviews in the manner of 'a series of friendly conversations' (Spradley 1979, 58). Most of our time was spent at three startups and the innovation incubator *InnoCubator*, all of them based in Bangalore. The startups were, in a sense, referents of nascent enterprises challenging the standard notions of scale. Apart from *Chalk Test*, this included a home and office maintenance platform which engaged with local electricians, carpenters, masons, and artisans in an ethical manner, which we will call *Nuedle*, as well as a bespoke vernacular language learning and translation services app which employed women from socially and economically challenged backgrounds, which we refer to as *Diverstics*. The fact that all three were within the purview of the project on scale which *InnoCubator* had commissioned additionally validated their potential as exemplars of entrepreneurship negotiating dominant scalar narratives.

To establish the subjectivities of scale imposed upon entrepreneurs, and also garner a sense of the alternate Weberian archetypes of entrepreneurs as opposed to those frequently projected and reinforced in public culture, we conducted semi-structured interviews with academics, other entrepreneurs who had realised scaling-up projects, investors, startup advisors, and staff from other established startup incubators and accelerators in not just Bangalore, but also the Delhi National Capital Region, as well as the cities of Mumbai, Kochi, Bhubaneswar, Chennai, and Hyderabad.

Having founded a bootstrap startup in Bangalore ourselves, a year prior to the study, helped serve as a phenomenological anchor during our ethnographic engagements. With due reflexive caution, we have alluded to our experiences in this paper, in the manner of 'embodied, intersubjective, temporally informed engagements in the world' (Desjarlais and Throop 2011, 92) rooted in the question regarding, and in confrontations with scale.

## THE DOMINANT DISCOURSE ON ENTREPRENEURIAL SCALE

'India's *startup economy has been booming. The last decade has seen significant activity on multiple fronts including the founding of new startups, amount of funding and number of investment rounds, influx of global investors and startups, development of regulatory infrastructure, global mergers and acquisitions, and internationalization. Entrepreneurial success stories abound. At last count, India had 26 unicorns, with eight new entrants joining the club in 2018 alone'* (Knowledge@Wharton 2019).

Over the meetings and tea-stall hanging out we did with the team of *Nuedle*, the one emotion we encountered time and again was that of anxiety, stress, and fatigue, all rolled into one. *Nuedle* had its origins in the challenges which the three founders had faced in getting everyday electrical and woodwork maintenance jobs done when they had moved to Bangalore over seven years ago. Bringing old fashioned relationship-building with local networks of electricians, carpenters, masons, and artisans, to a technology platform serving as a marketplace for home and office maintenance services, had led to early successes, unearthing (in the words of its founders) a 'big enough problem to solve for' and a 'huge market opportunity' which had 'delighted the investors' (Field Notes Extract 2019).

Figure 1 captures the responses to a survey administered by us over the summer of 2019, to *Nuedle's* three founders, its thirty-odd employees, a handful of its hundred-plus 'vocational professionals', as well as its advisors and investors. The survey asked the respondents to stack rank a list of 22 possible strategic focus areas for the firm, assuming a three-year horizon. Unbeknownst to the respondents, the areas had been categorized, as Processes, Financials and MIS, and Creating Value. The responses were normalised to identify the top five, bottom five and middle range of perceived priorities. Notwithstanding the relatively long time horizon, what is noteworthy from the table is that the top five priorities did not have any representation from the Creating Value category. This remarkable finding was perhaps duly qualified by the *Nuedle* founders' callouts, which we noted over the course of our fieldwork, such as being 'in a reactive, stressful phase for a long time now', 'working on transactional, administrative and tactical activities' because 'processes are not in place', being 'far away from where we had planned to be in 2019', not having been able to 'scale up', and that 'funding is a major challenge' (Field Notes Extract 2019).

The example of *Nuedle* demonstrates how conventional, dominant narratives of scale had come to act in the everyday operations of startups in the manner of Foucauldian power/knowledge subjectivities, thereby shaping their imaginaries. Yet even then, and as our research shows, these discourses were seldom unidimensional. A review of journalistic scholarship, thought papers, and reports on the startup ecosystem in India at the conjuncture of our study, highlights the emerging, negotiated and multidimensional aspects of scale, even if within the dominant frames, as we critically interrogated scalability and scaling-up as lived, socio-cultural experiences on one hand, and studied the assemblages which have contributed to its realisation on the other hand. Rooted in ethnographic enquiry and a review of public culture, we now turn to these narratives as resident within the dominant discourse.
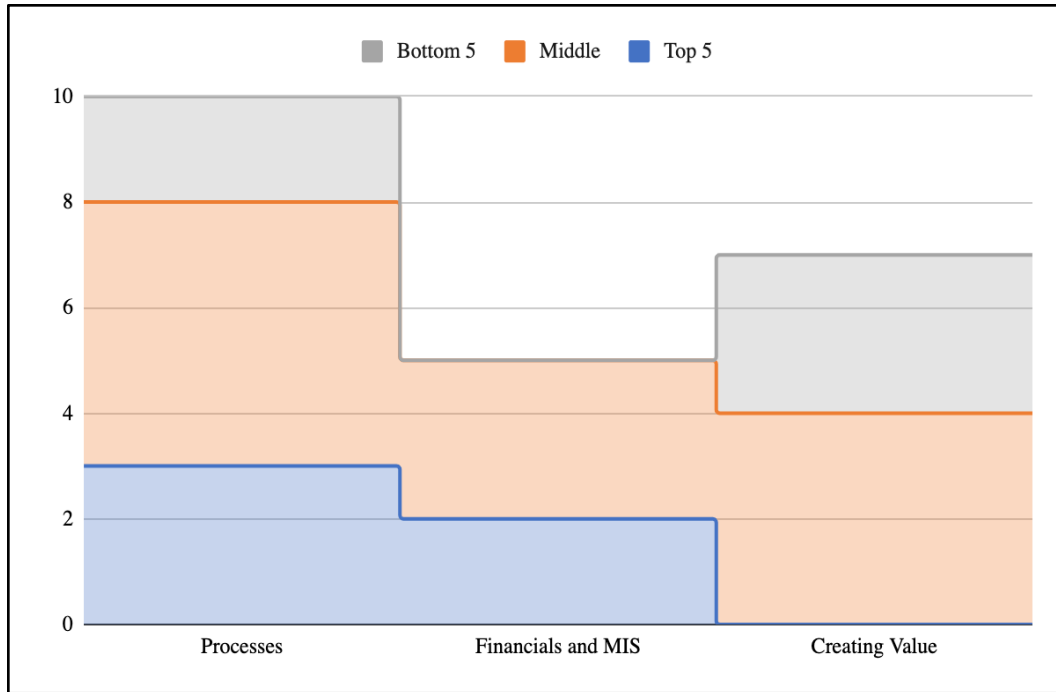
Figure 1: Survey responses at *Nuedle*

## The Successful Entrepreneur

Our research suggests that narratives in public culture privileged a certain archetype of a successful entrepreneur, even if they did not disregard the multiplicity of their skills and experiences. Distinct social identities are incorporated in the archetype, thereby defining scalability as a factor of the entrepreneur's educational degree, prior work experience, domain expertise, and even age. Thus, a startup which is likely to scale has been normalised and understood as one which is 'led by a group of well-educated co-founders with several years, if not decades, of work experience between them' (Chitnis 2018).

Additionally, the *maturity* of a given startup ecosystem is described in the form of the rising incidence of fluidity among roles, entrepreneurs-turned-serial entrepreneurs, serial entrepreneurs-turning-angel investors, investors-turned-entrepreneurs and employees-turning-entrepreneurs. These sequences are simultaneously suggestive of another dimension of scale, that is, of *successful exits*. The evidence for a startup's scalability is being read as a factor of there being individual entrepreneurs who are likely to 'produce more successful exits … [for investors and founders] ... with their tactical, experiential knowledge and easier accessibility' (NASSCOM Zinnov 2019). Of note is the singular absence of the word 'exit' from the 2018 edition of the same report, making its debut only subsequently. Furthermore, this potentiality of successful exits has also brought more players into the ecosystem, as for example, global and corporate investors.

While the ecosystem has witnessed existing players gaining expertise and moving across roles, it is also beginning to stress the need for building greater diversity within its fold, specifically among first time entrepreneurs. For example, dedicated funds and incubators for women entrepreneurs have been established. We noted however, that these distinct players did not bring about any change in the ecosystem as such, but incorporated their efforts within the existing structures and narratives. This is tantamount to a *non-becoming* of scale, which does not solve for inherited social inequalities, such as lower representation of women in corporate leadership roles. In turn, this has led to its own implications, as encapsulated below:

'Women *are missing in the Startup India initiative because many women, who start their initiatives, are not in the limelight or mentored professionally. Additionally, when it comes to funding, women are not only scrutinized about how they'd manage their businesses, but also their families in parallel, which isn't a filter men are put through … Thus women need to break through filters to raise capital and grow their businesses'* (Saxena as quoted in Sharma 2020).

Another example of such a *non-becoming*, was evident from a discussion earlier this year on women-entrepreneur-focused platforms in a WhatsApp group for entrepreneurs we are part of. One of the members highlighted how endemic gender inequalities are extended to the world of startups, as he shared how at the time both his wife and he left their corporate jobs to pursue entrepreneurship, he was questioned on whether he had done a proper risk assessment while nobody posed this question to his wife. As he noted, many believe that 'for women … entrepreneurship is *ghar+*' (or *home+*), thereby implying that women are seen as entrepreneurs *only after* domestic responsibilities have been taken care of (Field Notes Extract 2020). Continuing the thread, another women entrepreneur, who had chosen to set up her own enterprise independent of her family business, shared that '..being taken seriously [is a challenge] that resonates' and that 'there are days when I think being a businesswoman … is totally pointless … but then the next day is a new day again' (Field Notes Extract 2020). In this manner, gender as a social identity has been *re-territorialized* in a Deluezian sense, as both an enabler for entering at the lower end of the continuum of scale, as well as a barrier, for progressing up the spectrum

## Technology as an Actant

Finally, our study points at how *technology* as a non-human actor or actant in the assemblage of the entrepreneurial project of scale has also been 'both changed by … circulation and changes the collective through …[its]... circulation' (Sayes 2013). Each wave of the startup ecosystem brings about the demand for new business models and the technologies which facilitate them. And while technology itself scales up in the form of deep-tech and advanced-tech-led solutions, it also demands a related scaling by human actors. Each business idea is now pitched as a tech-led business, and technology is hardcoded into the eligibility criteria for a startup to receive policy or funding support. Technology is also incorporated within the themes or areas that solutions are invited for. That said, in having become a go-to-actant for scalability, technology has also created entry-barriers of digital inclusion bordering on the tautological, which entrepreneurs without prior knowledge or experience in technology either struggle to surpass or attempt to negotiate creatively, as in the case of the founder-proprietor of a Bangalore-based training services company we spoke with, who positioned the firm as an 'innovation and design thinking

organisation' and advertised its social media followers as an attestation of its 'commitment to and focus on technology' (Field Notes Extract 2019). An idea without technology is thus relegated to the realms of non-scalability.

## Scale Begets Scale

A review of the government's policy documents on entrepreneurship and startups, as well as the calls for funding published on its websites reveals the protocols and processes adopted to demarcate measures of success, define eligibility and outline criteria for entry and exit along the continuum of scale. The Karnataka Startup Policy 2015-2020 established three qualifiers for an entity to be categorised as a startup. It needed to be technology-based, registered or incorporated in the state of Karnataka (especially if it was beyond being an early-stage startup), and employing at least 50% of its workforce, not including employees on temporary contracts, in the state of Karnataka itself. The exit criteria called out revenue (in the sense of crossing 500 million Indian Rupees), and age (as crossing four years since registration or incorporation). The NASSCOM Zinnov 2018 report classified an Indian startup as having been incorporated in 2013 or later, having founders of Indian origin, with product development being carried out primarily in India. The startup was required to at least have a prototype or minimum viable product (MVP) in place. In this manner, the criteria reinforced the distinct social markers for the actors, in the sense of nationality and catchment area for constituting the workforce, as well as other dimensions of *time*, such as the years since incorporation or registration and stage of development and funding, and space, through the need for the startup to have been incorporated or having its value adding functions within a geographical boundary. With each subsequent stage of scaling-up, the necessary evidence was acquired as at once a qualification and an exercise in preparing for the forthcoming stage. As the entrepreneur-founder of an established communications startup we interviewed explained, what 'every incubator, investor or founders' collective' will expect is for the founder to 'plan for growing the business step by step, market by market, product by product' and in the process, 'think ahead for and *anticipate* the team, the service and product offerings, and the market plans you need for the next phase' (Field Notes Extract 2019). In other words, the dominant, conventional practice of entrepreneurial scale can then be analysed, as at once 'the product of a dialectical relationship between a situation and a habitus' (Bourdieu 1977 [1972], 261), where entrepreneurs' cultural, social, and symbolic capital in the form of community and professional networks reinforces the inference that scale *begets* scale.

Our study also suggests that the startup ecosystem has begun to drive a level of consistency in its *language*, with actors across the spectrum presenting their asks for funding, pitches, and success stories within the conceptual categories of age, company stage, funds already raised, education, prior startup experience, and awards as evidence of external validation. In this Goffmanesque presentation of the entrepreneurial self, scale has found another ontological materialization in the shape of hashtags such as *#30under30* and *#40under40*.

## Possibilities Beyond the Dominant Discourse

Within the extant narratives, the ability to effectively jump ahead, through space and time, is also celebrated. In 2018, a relatively unknown B2B marketplace progressed to Series C funding of USD 225 million within a record 26 months. Similarly, the transnational discourses of *global ambitions* were underlined, being defined in terms of expansion to markets outside India, entry of global startups in India, as well as the formalization of international startup exchange missions to promote global flows of capital, products, and entrepreneurs. Scale evolved to be expressed as not only higher, but also faster and wider. In 2019, one-fifth of the startups in India were focusing on global markets. At the same time, national projects such as AADHAR (providing a 12-digit unique identity number to Indian residents or passport holders, based on their biometric and demographic data) fuelled the digital infrastructure and materialization of data in the form of India Stack. The target addressable market in India was accelerating as well and, in 2019, 47% startups were reported to be serving low and middle income groups (NASSCOM Zinnov 2019). This construct of leapfrogging was extended to the agendas of development as well, in the manner of social impact and scale going hand in hand. Social impact incubators and accelerators for example, now evaluate startups applying for their programmes, on the strength of being able to demonstrate scalability across the country as well as evidence a customer base as a display of commitment to the solution.

In other words, and as we have shown here, even the dominant and conventional notions of scale were beginning to demonstrate multidimensionalities. In turn, this points at questions in conflation with *InnoCubator's* quest to identify alternate models of value that could help support its incubatees in their journey of scaling, asking whether scale always moved in the same direction. Does scale, with its multidimensional ontology, always progress up the continuum? Could there not be an opportunity to maintain or augment impact while operating at the lower end of continuum, at a given scale? Could this suggest an alternative model of realising value? Was the focus to enable startups to scale up, creating a vacuum at the lower end of the scale? Were there entrepreneurs who were already surviving and even thriving in this space? Could the project of scaling up continue without tethering value at the lower points of the scale continuum? These are the questions we now turn to.

## TETHERING AT DIFFERENT POINTS OF THE SCALAR CONTINUUM AS AN ALTERNATIVE MODEL OF ENTREPRENEURSHIP

'When *I joined about two and half years ago, I was the first employee,'* said Aravind, the marketing lead at *Diverstics. 'Today we have twelve staff members, an extended team of over two hundred and fifty women who we call our Language Experts, and nearly 1.5 million downloads and installations of our app. Yet we continue to remain true to our vision of using technology to help people learn vernacular languages for everyday use.' 'The women, our Language Experts… they are very dear to us,'* added Arpita, the founder-entrepreneur. '*We have not met even half of them in person, what with some living in villages and towns as far away as the eastern states of Bengal and Odisha. They are the ones who help in both creating and validating vernacular content on the platform, and that is a key part of our value proposition, in the sense that content is easy to follow and, if I can dare say so, colloquial. I mean at the end of the day, Diverstics is all about the spoken language, the everyday language. So what these women do,*

*essentially comprises our core competitive advantage. But what we truly treasure is the value which we have been able to create for these women. By providing them with incomes, we have enabled financial independence, and thereby, secured empowerment. As a woman and as a human being, that is very important to me.'* 'This deep respect and empathy for different members of our ecosystem, not just our app users, is intrinsic to our culture,' asserted Aravind. 'And we have sustained it by separating out the low-effort, continuous revenue generating business of language translation services. That keeps the engine running, and our investors are okay with our financials.' 'At the same time, we are working on stabilizing our technology platform, so that we can leverage the first mover advantage we currently enjoy in this space,'* qualified Kiran, *Diverstics'* technology lead. *'After all, we need to be prepared for a time when others jump in. Even if the likes of Reliance Jio or Google were to make an entry, our technology, data, and the service and quality levels they can deliver, should be able to compete against their scale.'*

Diverstics' office in South Bangalore was spartan. The decor was almost brutal, the furniture and equipment but utilitarian. Yet from the appreciative clapping all around the meeting room at Kiran's words, it was evident that the antiseptic and functional materiality of the office was not mirrored by the clearly enthusiastic and closely knit team. We could not help but ask as to what they believed was needed to be done in preparation for eventual competition. *'We need to stay on top of the right set of performance metrics,'* Arpita answered confidently. *'The number of installed cases, our monthly active users, user retention rates, and revenues from the B2C product and B2B services businesses taken separately are some of the measures we track at the moment. We need to be certain of what we are doing, and when we do it. I mean, we need to grow for sure, but we cannot lose our grip on the business we have built, or the culture which we have so painstakingly nurtured. There will always be the next opportunity. And we can be big. But we cannot afford to grow too fast. Everyone does not have to be a Jack Ma, Mark Zuckerberg, or Jeff Bezos. We need to be extremely careful and focused on who we bring on board, whether investor, staff, or extended team.'* (Field Notes Extract 2019)

Diverstics' case proffers itself as an exemplar of effectuation (Sarasvathy 2001) in the sense of the entrepreneur not operating solely under the boundaries of the intended effects of scale, such as expansion and grabbing market share, but instead creating the market as a potentiality of existing means, that is, 'through a network of partnerships and precommitments' (Burt, quoted in Sarasvathy 2001, 254). Thus, impact in the sense of women empowerment is consciously arranged alongside investor-friendly metrics such as revenue, users, and new markets, in an inter-scalar reinforcement of what could be its area of focus, given the stage it is in. *Diverstics'* focus on technology which is at once accommodative of expansion as also a catalysing of a stable business model, is anchored in Philips' situated assertion of 'a totalizing coherence of the overlapping scalar dimensions, a mutual propping up of each other' (Carr and Lempert 2016, 14).

Thus, *Diverstics* represents but a small, albeit growing, breed of entrepreneurs in India who are seeking to redefine startup scale in the inter-scalar sense of being both multidimensional as well as temporally situated. In this sense, they contest the dominant narratives of a unidirectional trajectory with future imaginaries where, as we have explained before, scale begets scale. Instead, they choose to lay anchor in and explore the potentialities which their temporal-space-present offers: they choose to be *tethered*.

Perhaps our own example demonstrates this better. In 2018, we co-founded a design research firm which sought to bring ethnography as a praxis to the field of management consulting. Our earliest projects were largely won on the strength of our professional

networks. Thereafter however, as word of our work spread, we were able to establish a run rate of a project every month, sometimes more. By the time we were a fifteen-month old organization we had scaled up in every sense, with eighteen projects across India, the Middle East, and Scandinavia, as well as having gone from four founders to a team of ten staff members. Yet that is precisely when we felt it was necessary to consolidate. We invested in specific training for the team, doubled down on establishing a distinctive, and researched client methodology. That was also the time that the two of us took time to pursue a second degree, this time in social anthropology, to add academic teeth to our offerings. In this way, we consciously chose to remain tethered, at the point of scale which we were.

In her paper titled Meaningful Innovation: Ethnographic Potential in the Startup and Venture Capital Spheres, Haines (2016, 175) posits that 'diffusion' or scale tends to constrain the potential of the technology startup ecosystem to contribute to 'more powerful innovations', since 'truly disruptive innovations ... by definition take time to diffuse'. She highlights how funding is precluded by evidence of scale, as she advances:

> The focus, rather, is on evaluating whether the product will scale before actually fully developing it. The process moves from finding potential early-adopter customers for an idea, to refining that idea based on how they may use the product, to then developing the actual product. The potential for diffusion precedes the innovation. (Haines 2016, 178)

Haines' powerful pronouncements are mirrored in our field observations. There is the founder of a grassroots organization which has been working to gainfully channel corporate social responsibility funds across South Asia, who tables that 'everything needs to run its its own due course' as he narrates how it took almost fifteen years from the time that he had first spoken with its chairperson, for his firm to start working with a regional enablement organisation, as 'scale and replicability, which go hand-in-hand, take time' (Field Notes Extract 2019). Equally remarkable are the founder-directors of a big data analytics company which has been bootstrapping since 2012 even as it has grown to operations in India and Singapore, who voice their belief in staying rooted in the core offering they are bringing to the market, instead of pursuing the 'glamour of wider markets, whether customers, geographies or product lines' (Field Notes Extract 2019). And finally, there is the affirmation by the now-profitable augmented reality start-up who says that, given an opportunity, he would go back in time to not raise the investments he then had, as he laments that while they certainly benefited from the money having come in, the '10x problems also led to 10x problems' as the 'fundamental drive of the investor is to scale, and find a buyer notwithstanding the actual work' (Field Notes Extract 2019).

In short, our research confirms that notwithstanding the more dominant narratives and discourses outlining scale as 'taking wings' or 'achieving escape velocity' (Startup India 2020), often aligning themselves with a 'mechanism of panopticism' (Foucault 1991, 216) which investors privilege by focusing on scale as a reflection on the entrepreneur's performance and capabilities, startups are increasingly beginning to explore potentialities and innovation at the existing point of the scalar continuum they find themselves on, through an exercise in consolidation and *tethering*. Furthermore, and running counter to popular perception, tethering as a means of seeking an alternative approach to value, is emerging in

entrepreneurship across sectors and models, both those which are for-profit as also those categorised as social enterprise.

What this does ask then, is as to what tethering at different points of the scalar continuum could offer, as an alternative model of entrepreneurship for the startup ecosystem. A possible answer is the opportunity for building relativity of scale. There is for example, the case of an established for-profit enterprise, which equips each of its clients, an overwhelming majority of which are for-profit sole proprietorships or limited liability partnerships, with the wherewithal to realise a steady earning potential grounded in the socio-economic *mise en scène* to which its founders, staff, and customers belong. Yet as our research indicated, it is not just the startup enterprises themselves, but even the institutional mechanisms and frameworks which are beginning to not just support but also facilitate such alternative models of value. Thus, alongside its efforts to set up incubation centres with a view to helping innovative startups become 'scalable and sustainable enterprises', the Atal Innovation Mission (AIM) as the Government of India's initiative to 'promote a culture of innovation and entrepreneurship in the country' has also begun to set up Atal Community Innovation Centres (ACICs) to 'focus on underserved/ unserved regions', it finalises plans to launch Atal Research and Innovation for Small Enterprises (ARISE). In fact, ACICs are part of a growing breed of incubators which are looking at incubating startups from low-income towns and districts which are solving for the *underserved and underprivileged.*

In drawing this chapter to a close, we offer a final ethnographic example of such scalar relativity. As part of our project, we carried out an ideation exercise for and with the founders of *Chalk Test.* Participating in the process were the startups portfolio managers from the incubator-investor *InnoCubator*, two of the startup's industry mentors, as well as members of a research team from *InnoCubor* which provided technical assistance to startups. For an exercise which started with the problem statement of scalar connotations, reading as 'how might we clarify our value proposition(s) for different users so as to have a commercially viable offering', it was remarkable that the principal areas of action were identified around positioning the app for only one set of stakeholders instead of all categories of users on one hand, and also working on strengthening but two identified internal processes (Field Notes Extract 2019). And what made this incident all the more noteworthy was the additional investment by *InnoCubator* in *Chalk Test*, which followed a few months thereafter, on the strength of what was effectively an exercise in tethering or consolidation, at the point which the startup was. Thus in tethering, it is not just that there are rising instances of scale's traditional, even hegemonic narrative, being resisted. Rather, and perhaps more significantly, the discourse is increasingly being negotiated in the form of its hitherto *nonscalable* articulations, to borrow from Tsing (2016), targeting scale at that point. Scale which is at once, tethered.

## TOWARDS AN ONTOLOGY OF DESIRE, AS A WAY IN AND WAY OUT

'Duniya *badalne ke sab ke apne-apne tarikey hai (Everyone changes the world in different ways)'* – from the Netflix movie *Upstarts* (Pawar 2019).

We explored an ethnography of mass media to reflect on how the imagined community of startups in Bangalore is depicted and consumed by audiences. The movie *Upstarts*, released on the OTT media service Netflix in 2019, was touted by its director, Udai Singh

Pawar, as the answer to yet another transnational discourse. 'Why is there no Indian equivalent to international film "The Social Network"?' (Outlook India 2019). Diffusion or scale does indeed take the shape of myriad concomitant flows.

Pawar drew upon his own lived experiences and interactions to advance an emic perspective, 'I studied at IIT Kanpur, and worked at Microsoft Research for three years. I have a background in Bengaluru because I lived there' (Outlook India 2019). The feature film follows three young men, from 'small town India', in yet another reiteration of the startup founder archetype. Tragedy begets resolve, as they are inspired to help provide the underserved with access to medicines. They design a technology led logistics solution which is essentially an app, register their own company, and start pitching their idea to funders as the *natural next step*. An inability to portray a *clear value proposition* leads to many failed pitches, before a chance encounter with an angel investor at an airport turns fortuitous. Money flows in, and then seamlessly, scale becomes a reality. The film plays on familiar tropes and establishes binaries as points of conflict: impact versus scale, founders versus investors, male privilege versus women entrepreneurs, and even interrogations such as 'what has greater value, their dreams or their friendship?' (Outlook India 2019). Interestingly, any desire to do something different is shown to exist outside of the project of funding and successful scaling. Two co-founders having an ideological conflict with the investor opt out after a certain progression of scale, as a result of dissonance with their desire for creating 'values that make for positive change' (Haines 2016, 197). The remaining co-founder then assumes the mantle of CEO, is hand-held by the investor and further progresses along the continuum. In this manner and form, the anchoring values are relegated to the background, and eventually abandoned at the behest of the funders. It makes an insignificant ontological appearance as a small team inconspicuously alluded to as the *NGO department*, that is clearly demarcated from the rest of the firm, in terms of space (separate workspace) and time (not keeping pace with the growth). A woman entrepreneur, who is a friend of the principal protagonist (the founder-CEO who continues with the firm and the investors), struggles to get funding (and screen time) for her own endeavour of 'creating a mental health support system', since it is not viewed as being investment-friendly. In addition, this venture is even positioned as *non-tech* with the women entrepreneur shown as personally engaging with an individual to save him from taking his own life. Finally, the founder-CEO is dismissed by the board for *non performance*, and utilises the opportunity to anchor back to his *desires*, reestablishes communitas with his two estranged co-founders, and is reintegrated with his values as he goes to working with his NGO team in a village on real issues. The social drama is resolved only when scale is sacrificed for values and societal impact. In a media interview, the director, Pawar asserted that '...the film is realistic and authentic, and based on hundreds of true stories' (Outlook India 2019). This claim entrenches the popular understanding of the startup ecosystem in Bangalore.

In this manner, we see that the film outlines desire as a way out of the structure of scale, and never as an instrument of negotiation. This notion finds resonance in Haines' research as well. Taking the example of an erstwhile startup, *Obatech*, in Indonesia, Haines outlines how Venture Capitalists, privileging short term returns over long term vision, force 'a distinctive shift in values—a shift that moves teams from doing something potentially meaningful and of value for a particular type of end user to doing something that potentially leads to value for the VC firm' (Haines 2016, 195). She builds the case for research in the ecosystem to examine emerging 'domains of interest' and 'to explore such contexts and

routines and identify areas of opportunity … as areas for positive disruption' (Haines 2016, 192). She puts forth compelling evidence for the relevance of anthropological theories and ethnographic approaches, in the manner of enabling a better framing, understanding, and assessment of teams and funding decisions (by the Venture Capitalists) as well as embedding value in the innovation process.

Our research however points us to another set of realities or multiplicities. We borrow an epistemological concept from Ravi Sundaram's study of pirate culture in Delhi (2010) to explore this. Can the entrepreneurial desire for value be analysed as the 'contagion of the [un]ordinary' (Sundaram 2010, 15) that does not resist the discourses of power, but revels in its cracks and gaps? Should concomitant streams of nonscalability be viewed as lines of flight which, even while running along the periphery of incumbent structures, always carry the possibility to escape and thus, create potentialities for creativity and disruption? Our research shows that the lines of flight of desire can enter, exit, or run along the continuum of scale at will. Startups, while focused on vision and impact, enter projects of scaling to be able to experiment and pivot. This is aided by the fact that the system is inherently characterized by a high tolerance for failure. 'Entrepreneurs cum investors, however, tend to have a higher risk tolerance, which in theory helps spur innovation at a much higher rate than corporate R&D' (Kaplan and Lerner as quoted in Haines 2016, 188). As one Venture Capitalist shared, 'Sometimes, funding enables that pivoting and subsequent acceleration' (Field Notes Extract 2019). At the same time, one cannot escape from the focus on conventional performance metrics endorsed by Venture Capitalists including retention, growth, acquisition growth, daily and monthly active users, lifetime value of a user, acquisition cost, profit margin, and potential market size, amongst others (Haines 2016, 193). While managing these metrics as feeds for funders, investors, incubators and accelerators, the founders also retained their focus on sustenance of value. The multiplicity does not end there. As we have mentioned earlier, there were entrepreneurs who chose to remain entirely outside of these structures or 'bootstrapping' for 'keeping our head above water for a long time may be sufficient' (Field Notes Extract 2019). And then there were others who entered the ecosystem, but exited after realising that untethered growth was unsustainable, 'we had 10x money and 10x problems, and chose to scale down' (Field Notes Extract 2019) and chose to *assemble* again, 'And that's when we doubled on efficiency - and it was this reason that we are profitable today... core business logic cannot be done away with' (Field notes Extract 2019). Our research opened up a future line of enquiry. It was evident that entrepreneurs have started finding meaning in alternative modes of value, which can even exist alongside projects of scale. It would be interesting to see if and how these desires as lines of flight in turn facilitate a Deleuzian *becoming* of the startup ecosystem. This presents another opportunity for ethnographic research as a method to study emergent trends in the reterritorialization of the startup ecosystem.

## CONCLUSION

In a world at once digitally connected and dispersed, Horst (2012, 72) reminds us of the need to remain committed to the 'classic anthropological ways of knowing', including but not limited to an 'attention to change over time'. The suggestion's singular message exhorts us to keep on returning to the field, an act rendered all that more germane for us, being as we are both ethnographers and entrepreneurs. And we have faithfully returned since the

time of our research in 2019, to note developments in keeping with our observations, which confirm our ethnographic readings. Scale as facilitated by language, for example, is beginning to be challenged through incubators which are now looking to provide entrepreneurial support and guidance in the vernacular, not just in English. Their argument is built on the view that a restricting of communication, incubation support, and mentorship to the English language alone, is an effective limiting of the pool of entrepreneurs, not too different from Anderson's imagined communities premised on a 'consciousness' imagined through the 'unified fields of exchange and communication' (1991, 44). There are also drives to onboard an increasing number of women entrepreneurs, mentors, and coaches, in an attempt to bridge the gender gap across the continuum, often in conjunction with a focus on a regional startup ecosystem. An example is the 'Her&Now' programme being led by the Government of India's Ministry of Skill Development and Entrepreneurship (MoSDE), along with the Deutsche Gesellschaft für Internationale Zusammenarbeit (GIZ) GmbH, to better the 'framework conditions for businesses managed by women in India' in the states of northeastern India, as well as those of Rajasthan and Telangana (GIZ 2020). As an initiative which focuses on historically neglected regions, there is an institutional push for *tethering*.

In summary, there is no doubt then, that the question of scale is evolving. Dominant discourses of power/knowledge remain, but alternative perspectives have started to emerge alongside, in a decidedly syncretic manner of assemblages. The actors include not only startups themselves, but institutional enablers such as incubators and accelerators, as well as regulatory and administrative machinery. Even more significantly, and situated within this multidimensionality of scale, are projects of scale-making both tethered or at a point, as well as in the form of potentialities of desire comprising alternative ontologies of value. And while our ethnography was situated in India's startup capital, there is much to learn for stakeholders of startup ecosystems the world over, whether entrepreneurs, funders, incubators, accelerators or governments and regulators. For there is no shame in not scaling. It is no longer construed as failure or a rite of passage for an entrepreneur. Everybody then, is a winner.

**Gitika Saksena** is a Director at LagomWorks Consulting. LagomWorks is a research and innovation consulting firm, leveraging principles and methods of Design Thinking and Ethnographic Research to drive transformation and innovation. Prior to LagomWorks, she was a Vice President at Accenture Technology in India, where she led the strategy for various talent initiatives. She has just completed her second Masters' degree in Social Anthropology from SOAS University of London on a British Chevening Scholarship.

**Abhishek Mohanty** is a Director at LagomWorks Consulting. LagomWorks is a research and innovation consulting firm, leveraging principles and methods of Design Thinking and Ethnographic Research to drive transformation and innovation. Prior to LagomWorks, he was an Associate Director in Management Consulting at PwC India. Abhishek has just completed his second Masters' degree in Social Anthropology from SOAS University of London on a British Chevening Scholarship.

## REFERENCES CITED

Anderson, Benedict. 1991. *Imagined Communities: Reflections on the Origins and Spread of Nationalism.* London: Verso.

Bourdieu, Pierre. 1977 [1972]. *Outline of a Theory of Practice.* Cambridge: Cambridge University Press.

Briody, Elizabeth K., and Alex Stewart. 2019. "Entrepreneurship: A Challenging, Fruitful Domain For Ethnography". *Journal Of Business Anthropology* 8 (2): 141-166. doi:10.22439/jba.v8i2.5846.

Carr, E. Summerson, and Michael Lempert, eds. 2016. *Scale: Discourse and Dimensions of Social Life.* Oakland: University of California Press.

Chitnis, Shailesh. 2018. "The Secret Sauce Behind A Successful Indian Start-Up". *Mint*, January 11. Accessed May 8, 2020. https://www.livemint.com/Companies /gIzjSNuBykoPFmPIVopk5J/The-secret-sauce-behind-a-successful-Indian-startup.html.

Desjarlais, Robert, and C. Jason Throop. 2011. "Phenomenological Approaches In Anthropology". *Annual Review Of Anthropology* 40 (1): 87-102. doi:10.1146/annurev-anthro-092010-153345.

EPIC. 2020. "Theme." Accessed May 31. https://2020.epicpeople.org/theme/.

Foucault, Michel. 1991. *Discipline and Punish: The Birth of the Prison* (translated by Alan Sheridan). New York: Vintage Books.

GIZ. 2020. "Empowering women to become entrepreneurs." Accessed August 3, 2020. https://www.giz.de/en/worldwide/75649.html.

Goffman, Erving. 1971. *The presentation of self in everyday life.* Harmondsworth: Penguin Books.

Haines, Julia Katherine. 2016. "Meaningful Innovation: Ethnographic Potential in the Startup and Venture Capital Sphere." *Ethnographic Praxis in Industry Conference Proceedings 2016*: 175-200. Accessed May 18, 2020. https://www.epicpeople.org/meaningful-innovation/.

Horst, Heather A. 2012. "New Media Technologies in Everyday Life." In *Digital Anthropology*, edited by Heather A. Horst and Daniel Miller, 61-79. New York: Berg.

Knowledge@Wharton. 2019. "Three Waves: Tracking the Evolution of India's Startups." Accessed July 30, 2020. https://knowledge.wharton.upenn.edu/article/three-waves-tracking-evolution-indias-startups/.

Mansfield, Nick. 2000. *Subjectivity: Theories of the Self from Freud to Haraway.* New York: New York University Press.

NASSCOM Zinnov. 2018. *Indian Tech Start-Up Ecosystem: Approaching Escape Velocity.* Noida: NASSCOM India.

NASSCOM Zinnov. 2019. *Indian Tech Start-Up Ecosystem - Leading Tech in the 20s.* Noida: NASSCOM India.

Outlook India. 2019. "Upstarts: The Story Behind India's First Film On Startups." *Outlook*, October 17. Accessed August 3, 2020. https://www.outlookindia.com/newsscroll/upstarts-the-story-behind-indias-first-film-on-startups/1642485.

Pawar, Udai Singh, dir. *Upstarts*. Bandra West Pictures, 2019. https://www.netflix.com/title/80998890.

Philips, Susan. 2016. "Balancing the Scales of Justice in Tonga." In *Scale: Discourse and Dimensions of Social Life*, edited by E. Summerson Carr and Michael Lempert, 112-132. Oakland: University of California Press.

Samudra, Jaida Kim. 2008. "Memory in our body: Thick participation and the translation of kinesthetic experience." *American Ethnologist* 35 (4): 665-681.

Sarasvathy, Saras. 2001. "Causation and effectuation: towards a theoretical shift from economic inevitability to entrepreneurial contingency." *Academy of Management Review* 26 (2): 243-288.

Sarukkai, Sundar. 1997. "The 'Other' in Anthropology and Philosophy." *Economic and Political Weekly* 32 (24): 1406-1409.

Sayes, Edwin. 2013. "Actor–Network Theory And Methodology: Just What Does It Mean To Say That Nonhumans Have Agency?" *Social Studies Of Science* 44 (1): 134-149. doi:10.1177/0306312713511867.

Sengupta, Vishnupriya and Suvarchala Narayanan. 2019. "India's New Unicorns: The world's largest democracy is becoming a seedbed for billion-dollar startups." *strategy+business*, October 21. Accessed July 28, 2020. https://www.strategy-business.com/feature/Indias-new-unicorns?gko=76038.

Sharma, Neetu Chandra. 2020. "Why women entrepreneurs are missing from India's startup story." *Mint*, February 21. Accessed July 30, 2020. https://www.livemint.com/companies/start-ups/why-women-entrepreneurs-are-missing-from-india-s-startup-story-11582267821850.html.

Spradley, James. 1979. *The Ethnographic Interview*. Fort Worth, Tex. ; London: Harcourt College Publishers.

Startup India. 2020. "Karnataka Startup Policy 2015-2020." startupindia.gov.in. Accessed July 29, 2020. https://www.startupindia.gov.in/content/dam/invest-india/Templates/public/state_startup_policies/Karnataka_Startup_Policy.pdf.

Sundaram, Ravi. 2010. *Pirate Modernity: Delhi's Media Urbanism*. Oxford, New York: Routledge.

Tsing, Anna Lowenhaupt. 2019. "On Nonscalability: The Living World Is Not Amenable to Precision-Nested Scales." *Common Knowledge* 25 (1-3): 143-162.

Variyar, Mugdha. 2018. "Seed Stage Funding Falls 20%, Late Stage Funding Grows 259%: Nasscom Startup Report." *The Economic Times*. October 25. Accessed July 29, 2020. https://economictimes.indiatimes.com/small-biz/startups/newsbuzz/seed-stage-funding-falls-20-late-stage-funding-grows-259-nasscom-startup-report/articleshow/66361259.cms.

**THEMATIC SESSION**

# Trees and Forests

When should systems know us as people and not as just data? Who watches the watcher when abstractions and aggregated data are used to make decisions about our lives? The flow of information and data between human and machine systems are a source of both progress and anxiety. In this session, we consider how a change in perspective (scale) leads to shifts in how data is contextualized and understood. Implications are presented on the ethnographer's role within these systems and as the watcher of them.

Session Curators: Afra Chen, Lisa Kleinman, Carry Yury

**PECHAKUCHA**

# How I Learned to Stop Worrying and Love Surveillance

SUSAN FAULKNER, *Intel Corporation*

*When a man rang our doorbell late at night and claimed that his teenage daughter was in our house, but she wasn't, my husband and I considered getting a doorbell cam. With camera surveillance and facial recognition becoming more commonplace, we wanted a privileged view of our surroundings, and a sense of control over what was happening on our doorstep. But, while we wanted the doorbell cam to see our late-night visitor if he ever came back, we knew it would also see us coming and going, and living our lives. We put the thought of a camera aside, but a few weeks later another uninvited guest knocked on our door. The coronavirus arrived in the US with a vengeance, and suddenly everybody we saw was a possible carrier of contagion. My husband and I, the people who had rejected a little doorbell cam as being too invasive of our privacy, started daydreaming about living in a country like Korea where our privacy and independence would be tested, but where our interdependence as humans would be understood. Where knowledge of everyone's comings and goings was a matter of life and death, and where we would know enough about the public health crisis around us to do something useful about it.*



"God Bless America" © Susan Faulkner

**Susan Faulkner** is a Senior Researcher and research manager in the Client Computing Group at Intel Corporation where her work has focused on remote education, the future of connectivity, media creation and consumption, the emergence of new interaction platforms including facial recognition, and the role technology plays in people's daily social transitions. She has an undergraduate degree in Comparative Literature from Brown University, and an MA in Documentary Filmmaking from Stanford University. susan.a.faulkner@intel.com

CATALYST

# Scale and the Gaze of a Machine

RICHARD BECKWITH, *Intel Labs*
JOHN SHERRY, *Intel Labs*

*Scale suffuses the work we do and, recently, has us considering an aspect of scale best suited to those with ethnographic training. We've been asked to help with scaling up one of the latest blockbusters in high tech – deep learning. Advances in deep learning have enabled technology to be programmed to not only see who we are by using facial ID systems and hear what we say by using natural language systems; machines are now even programmed to recognize what we do with vision-based activity recognition. However, machines often define the objects of their gaze at the wrong scale. Rather than "look for" people or objects, with deep learning, machines typically look for patterns at the smallest scale possible. In multiple projects, we've found that insights from anthropology are needed to inform both the scale and uses of these systems.*

*Keywords: Deep Learning, Human Scale, Ethnographic Insights*

## PEOPLE THINK AT A HUMAN SCALE

When we talk about "human scale", we refer to the sizes of objects and spans of time that people tend to think about. We humans don't *have* to think on the human scale. We *can* think on the scale of the universe or the atom. However, thinking at the human scale is natural; it is what allows us to collaborate; it allows us to see the reasons in another's acts; it supports our sociality. Although we can argue with an imposition of "rationality" on broad of swaths human thought (e.g., Malinowski 1922/1984), we also must admit that it is typically rather easy to attribute a rationale to what a person has done. We naturally "see" what other people are doing; machines do not.

Why don't machines just see like humans? Humans program the machines after all. The reason is that machines would need to be programed to see at a human scale and, at this point in time, that hasn't been the case. It's quite hard and there are alternatives. Machines have been programmed to a surprising level of accuracy, to be sure, but that's not enough. You can be accurate and yet not correct. The human ability to see what others are doing – this "vision" – is not the same as being able to describe the outward behavior that people have engaged in. The social sciences became convinced of that disconnect following the fall of behaviorism. Now, the social sciences rarely provide an "objective" description of the "behaviors" of others, rather, we offer what might be called a "preferred description." (Searle, 1983). Someone might describe another's behavior as alternating movement of the legs across a floor, but this would likely not match how the person would describe it themselves. An observer might say that a subject has walked to the north, which may be true, but the walker may not even have known the direction. It's more likely that the person being observed had thought that they were walking to the exit. "Walking to the exit", then, is the preferred description and these descriptions are easy for humans to generate about each other. It seems fairly obvious that a person watching that walker would say the same thing, and perhaps this is what Malinowski may have had in mind – that he could look at

Trobriand Islanders and their culture and imagine why they would travel great distances to bring some long-held possession to be held by another.  That attribution is thinking at a human scale (e.g., Dennett on the "intentional stance", 1978).  It's so much easier to collaborate with, to trust, another whom you can at least convince yourself that you can understand.  So, it can be a real problem when "thinking machines" don't think like us.

## Machines Don't Have to Think in Human Ways

One of the reasons that our technology company hires social science types is to help to design technologies such that they are better partners.  It used to be we were asked to help make purely responsive computers that would fit with people's lives. Now, the computer can take initiative (Console, et al. 2013) and fitting in is so much more significant.  New technologies promise to be more connected to their environment and better able to understand and interact with people in more natural ways. That promise is where the problems start.  It's frequently the case that "high technology" is designed in a decidedly non-human way and we're here to tell the choir that these machines can be harder to collaborate with and harder to trust than people.  In many ways, what we are trying to do in our work is to help to create technology that can truly participate at the human scale or to point out when machines are incapable of working with that way people.

We'll detail some examples from the technology literature and briefly describe some cases that we're working on, but before that, we'll lay out a technology domain to which we will restrict our focus, one that is not only salient these days but which also highlights the value of the social sciences for technology development, namely artificial neural networks or, more simply, "neural nets".

## NEURAL NETS

Neural nets are the "iron horse" of the 21st century.  OK, maybe "neural net" is just a similarly inapt metaphor.  Iron horses weren't remotely horses and neural nets aren't remotely brain-like.  Despite not being horses, railroads have been remarkably useful as a means of transport.  They deliver goods, simplify travel, and can be quite reliable.  Neural nets can be remarkably useful, too. As many people know, neural nets are terrific at finding pictures of cats (Le, et al., 2012).  Moreover, neural nets are driving significant innovation in the computing industry. They have enabled improved multimodal sense-making and understanding (Owens and Efros, 2018), automated speech recognition (Chan et al., 2016), and natural language processing (Vaswani et al., 2017) and, then, there's that near magic we see with computer vision; and, it goes well beyond cats (Krizhevsky et al., 2012).

## AlphaGo

For famous example, AlphaGo, which debuted in 2016, was built on a neural net that was programmed to play the game Go (Silver et al., 2016).  Go is a two-player game where players capture space by putting colored playing pieces on a game board.  At its debut, AlphaGo beat a world champion Go player in four out of five games.  This was a surprise to nearly everyone, including the AI community, because Go is considered much harder than

chess for a computer and computer scientists had worked for years on computer chess before being able to beat a human champion.

While there are lots of different aspects to the game and the program, we want to focus on just one aspect here – the Go board, its moves, and how AlphaGo sees them. First, let's consider how humans see Go. The Go board is a grid of lines that form 19 squares across, 19 squares deep, and has 361 intersections. (361=19x19) These intersections are where a player puts the playing pieces – "stones" – which are black for one player and white for another. Players take turns in placing one of their stones on an empty intersection. The goal of the game is for a player to build continuous walls of their stones around sections of the board such that their walls enclose more space than their opponent's. When a player puts down a stone, it is to either build a wall of their own pieces or block their opponent from building a larger enclosure. Any surrounded stones of an opponent are taken as prisoners. Each completed game takes about 250 turns. It will be relevant in the paragraph after the next to have noted here that, for humans, reading the current paragraph once or twice would allow a person unfamiliar with Go to not only play the game but also create a functional board with playing pieces.

Humans see a Go board as a 19x19 grid on which walls are built with stones. That's not how AlphaGo sees the game. AlphaGo sees the Go board as one long vector with a separate element for each of the 361 intersections. The training data for AlphaGo consist of game length sets of these vectors with each consecutive vector in the set representing each subsequent move in a game. As AlphaGo sees it, each move in the game is represented by a new vector that is different from the previous vector by one element (i.e., the new stone) or more if an opponent has been surrounded and their pieces taken as "prisoners". The bottom line is that people playing Go see the building of walls around sections of the playing surface; AlphaGo sees patterns in a series of vectors.

Before playing a game against a person, AlphaGo will look at, literally, millions of games to see what patterns emerge in the vectors. Once AlphaGo has seen millions of games that were played, it can figure out how to win. More specifically, AlphaGo can figure out which next step (i.e., which change in one element) is most likely to lead to a win and with each step chooses the move it believes will get it closer to a win. In order to learn to play at the level it played, AlphaGo needed to see millions of games that had been played. Interestingly, in order to play at all, AlphaGo would likely have needed to have access to nearly as many completed games. This requirement of seeing millions of games, it must be noted, is simply not true of humans who can learn the game quickly (as noted in a previous paragraph) and people are unlikely to ever encounter a million games in their lifetime let alone by the time they've played their first opponent.

## Feature Engineering

To be perfectly honest, almost none of that is central to the argument we want to make. What we care about most is that the two-dimensional 19x19 grid on the board on which a person sees walls, AlphaGo sees as a simple line with pieces of data about the state of each cell (black, white, or empty) which forms patterns with the state of the board in nearby lines. That AlphaGo sees the state of the board as linear is quite significant since a line can have no walls. AlphaGo simply finds patterns in the sequence of changes between the lines within a game.

One can imagine that engineers didn't have to spend much time figuring out that a vectorized representation of a two-dimensional board was going to be good enough. They still had a single variable for each intersection and only three different states of those 361 intersections. Noticing patterns across elements isn't likely to be outside the ken of an artificial neural network and, frankly, there isn't much else going on in the training data that the machine would need to notice or would be distracted by. The system only needs to know possible next steps and the likelihood that a change in arrangement on the board will lead to a winner. So, the feature engineering for AlphaGo would have been fairly simple. Nevertheless, feature engineering is an important part of any neural network or machine vision system and is nearly always much more complex than what we've seen with Go.

In fact, deciding which features to include in training a neural net can be quite difficult especially in areas like vision or language which so often seem magical. Because of this difficulty, engineers have discovered ways to allow a program to find its own features. This is called "automatic feature engineering". Despite the fact that automatic feature engineering has some fairly significant issues, in many ways, it is the magic of vision and language neural networks and underlies the ability to find so many cats. Yet, it can lead to a particularly pernicious type of problem – inferences based on spurious correlations.

Spurious correlation errors are one of the more significant side effects of automatic feature engineering. Obviously, spurious correlations are not just a problem for deep learning. People fall prey to spurious correlations, too. Consider for example, the recent conspiracy theory holding that 5G radio towers cause Covid-19. The best evidence that proponents have for this theory are geographic heatmaps showing that, in February and March of 2020, Covid hotspots and the then-current 5G deployments lined up quite well. The correlation between maps looked compelling, and without a more sensible explanation, 5G could seem like a reasonable-enough theory. The reason for maps lining up, according to experts, was that Covid was hitting urban areas hard and urban areas are also where 5G rolled out first. The correlations between Covid and 5G were spurious. What is important to note here is that we can see the sense of people's mistaken explanations – "the maps lined up so well"; there is a transparency to the error.

Often, transparency of errors isn't the case with deep learning. In fact, sometimes the errors generated with deep learning seem inexplicable. Research on attacks against deep learning systems can demonstrate how opaque the reasons for an error can be. For example, researchers have created patterned eye-glass frames that will fool a state-of-the-art facial recognition system created with automatic feature engineering (Sharif, et al. 2016). This system was trained to recognize different celebrities. The automaticity in the facial recognition system had the system look for pixel-level differences between a number of photos that were labeled with different celebrity names. As with the Go board, the system looked at each picture as a long vector. That is, photos were seen as a long line of pixels. In these digitized photos, the pixels are row after row of dots, each of which is one color, not unlike the Go board with its 19 rows of 19 columns and three states per element. Photos are just more complex than a Go board: more rows, more columns, and more states per element. Instead of Go's three states, the colors of a photo can include 100s of options or more. So, an image is, like the Go board, seen as a vector, but a much longer vector with much more varied contents.

The complexity of digitized images means that there is a greater chance of spurious correlations. The photos of celebrities offered spurious correlations aplenty. The

researchers in this study found that they could design a set of colorful eyeglass frames, each of which appeared to have a random design, but the design would match a pixel pattern associated with a particular celebrity. The researchers discovered patterns that would fool the vision system into believing that one person was another. For example, despite the fact that the system was excellent at recognizing photos of Reese Witherspoon, a picture of her was mistaken for Brad Pitt when she was pictured wearing the Brad Pitt glasses (or other celebrities when other glasses were used). [We suppose we should mention that to most people, these two celebrities don't look much alike.] Any person wearing the Brad Pitt glasses would look like Brad Pitt as far as the system was concerned. Brad Pitt was identified by the pattern of pixels in the eyeglass frames (there were certainly other "random" patterns of pixels that happened to be associated with Brad Pitt but those on the glasses were sufficient for identifying him.) Despite being state-of-the-art, the facial recognition system fell for a spurious correlation. However, unlike the similarity of the maps of 5G and Covid, the correlation that the system found between name and pixel pattern was not something that a person could ever see. The patterned glasses don't even remotely look like Brad Pitt or any of his features. The errors would make more sense if the researchers had deployed prosthetic chiseled chins to make someone look like Brad Pitt. People simply don't hypothesize identity of others based on random patterns in pixels.

## PROBLEMS FOR ETHNOGRAPHERS

So, now we've covered neural nets and feature engineering and the problem with spurious correlations and can now turn to projects we've worked on to highlight some of the issues that ethnographers are best able to deal with.

### Communication versus "Natural Language" Networks

One of the projects that we are now working on is a system that will use deep learning to translate between American Sign Language (ASL) and English. The idea is to find patterns in videos of people signing and relate those patterns to simultaneous English translations. The videos we are using sometimes have ASL translated to English and, other times, English translated to ASL. In all cases, these videos include ASL and English that are intended to express the same content. The goal is to have an "end-to-end" system that learns from videos of signing and an associated translated text of the spoken language used as a "label" for the signed content. Tens of thousands of these labeled videos are required for the system to begin to learn to translate.

Given that the system's input streams include raw video, it will not be surprising to hear that the system will be looking at the video as a sequence of vectorized images with the top left corner of the video being the first element in the vector and the bottom right pixel being the last. The features that the system will discover are like those of the celebrity ID system – in that they are patterns of pixels associated with some label.

A knowledgeable signer of ASL would look at the video and see a sequence of meaningful tokens (i.e., morphemes) composed of a set of language specific building blocks (i.e., phonemes) but the deep learning system has automatic feature engineering and is learning without seeing or knowing anything about phonemes or morphemes and, further, is not being programmed to acquire them. By focusing on pixels and patterns of pixels, the

system is far simpler to program.  By focusing on these language independent features (i.e., pixels), problems with spurious correlations are rife. Systems in the future will be able to learn morphemes and phonemes first and acquire the language with that "knowledge".  This is the only way to avoid the problem of "spuriosity".  But this is only the beginning; the problem with pixels goes further than that.

Ethnographers trained in microanalysis can say more about what a knowledgeable user of ASL sees or how a fluent signer would construct and understand meanings.  What microanalytic techniques brought to the study of communication was to show where relevant data had been ignored in trying to assign meaning:  The weight of conversation is not carried only by syntactically words; there are non-linguistic gestures, postures, and eye gaze (Birdwhistell 1970, Kendon 1967, Schegloff 1998).  There's intonation, pitch excursion, and volume.  Conversation even moves forward with what is not said (Watzlawick, 1967).  These all fly under the banner of "microanalysis". What microanalysis brought to the more strictly behavioral concerns of the time was a research program that asked what needed to be considered in the way that people construct and understand meaning when they communicate.  This methodology is associated with anthropology as much as communication theory; both areas study meaning and the technologies and techniques with which meaning is shared.  There can be no question that a fluid and facile interpreter will need to consider these cues nor that a system meant to interpret must also consider them.  A job for the ethnographer working with deep learning is discovering both the right level of analysis and an ontology that makes sense…and then advocating for them.

## The Interpretive Stance and Machine Vision Networks

Part of the magic of these deep learning systems is not only that they can work at all but also how well they work once they do (remember all those cat photos).  Part of the problem, is that when they make an error, it will not be an error that a person is likely to be able to understand.  That is, it won't fail in a human way and a person working with it is unlikely to be able to determine what data it considered and how it was analyzed while making an inference.  When the system offers a solution, a user may find it difficult to know that it has failed.  Simply put, when the errors are not on a human scale, it is difficult for a person to be able to correct it, to work with it.

## How Do You Work with Failure?

Arguably, effective translation is crucial, and errors could be life threatening.  However, it is also the case that, in an operational system, an ethnographer will have insured that conversational methods of correction would be in place.  Perhaps an example where the system performs as an autonomous tool would help to highlight the potential risk of our misunderstanding how a machine sees.  Here's another example from the tech literature.

The boffins have taken deep learning's most common machine vision training set (i.e., ImageNet (Deng, et al., 2009)), played with something quite like the Brad Pitt eyeglasses noted above, and come up with something diabolical (Athalye, et al., 2018).  While it doesn't include celebrity photos, ImageNet is a database of one million images of many different classes of objects.  This database is used by many deep learning practitioners to build systems that identify new images of the object types included in ImageNet (like turtles and

rifles).  In this case, the boffins trained up a network so that it had world-class performance in identifying the object categories.

One of the object categories that is relevant for this story is that of "rifle".  Rifle plays the role of Brad Pitt here.  What's interesting is that these researchers used the seemingly random pattern of dots/pixels associated with "rifle" and did something akin to what the other researchers did with the Brad Pitt pixels.  Instead of eyeglasses, they manipulated a view of a 3D toy turtle with this random dot pattern.  Then, they rotated the turtle and placed the dots such that from every angle, the toy turtle looked like a rifle to the network. To the human viewer, the coloring wound up looking a bit like turtle camouflage. So, instead of a person wearing colored frames on a pair of glasses and then looking like Brad Pitt, a toy turtle was misidentified as a rifle. One can imagine negative consequences that could follow from having a child bring such toy into a protected area…very negative consequences and the reason for the error would not be at all obvious to those protecting that area.   Because of the way pixel-based systems work, one would hope that a security detail would never rely on one.  (Of course, police departments do use deep learning-based machine vision already (Harris, 2019).)  Clearly, there's more for the ethnographer to do.

## Collaborating with a Deep Learning System

People expect that others, whether a person or a system, will see things as they do.  We can learn a lot about how we see things by thinking about how we live and work with others. It's really important that we communicate and agree on what things are.  If we see something, we expect that an intelligent other will see the same thing and call it by the same name.  If someone calls something by a name we know, we expect that thing to be what *we* would call that name.

Communicating people don't necessarily agree on everything but, at least where collaboration is concerned, we usually mean the same thing with words.  Formally speaking, ontologies do not have to be identical, merely sufficiently overlapping and with a method for finding and resolving difference, if need be (Ludwig, 2016).  Some remaining differences are fine as long as we understand what they are.

For example, the Kahluli of Papua New Guinea consider the male and female birds of paradise to be different species and this is entirely consistent with their knowledge that the two come together for breeding (Feld, 1982).  This isn't likely to be of much consequence when dealing with a Kahluli person.  If you want to see a male bird of paradise, you simply ask to see one.  It doesn't matter that it isn't considered the same type of bird as the female. In fact, this is kind of what much ethnography has always been about:  How should we understand others outside our group?  Classic Ethnography is rife with examples of ontologies that aren't shared.  Ethnographers can work with that and explain how to understand each other.

In this way, ethnography, like anthropology more generally, assumes a rejection of radical incommensurability. This rejection of incommensurability means that the concepts deployed by one entity (individual or collective) can be understood by another.  When someone discusses their family, say, they may have in mind a different set of people from what another might assume but, if the two share sufficient beliefs, they can discuss the boundaries of the concept of "family".  Foucault would have called this an episteme (1971) and Kuhn, a paradigm (1962).  What's important is that our categories are fluid and we can

work within and, to a great degree, between them.  Ethnography assumes a level of commensurability sufficient that someone could explain another in terms that are understood.

## Practically Incommensurable and Practically Inscrutable

Unlike the subjects of ethnographic work, systems created using deep learning are practically incommensurable because they are practically inscrutable.  That is, in practice, such systems work with very different concepts from the people who work with them and it will take a lot of work to get to a point where differences can be discovered and resolved.

*Incommensurability*

If you imagine that a system is observing as you would and "describing" those observations in terms that you would use, a deep learning system could easily be seen as an unreliable observer; however, nothing is further from the truth.  The system is quite a reliable observer; under the same conditions, it will come to the same conclusions.  It is the expectations of the naïve user that is a problem because the borders of the system's concepts are considerably different from our own (consider Brad Pitt or the rifle).  Still, how can you rely on someone who tells you things that you know are simply wrong?  How can you work with someone you don't understand and with whom you cannot negotiate a shared meaning?  It is only by coming to understand the other's constraints.  An example from our work might help.

We work with another project that uses machine vision.  This one watches factory workers.  The goal of this system is to improve safety while, at the same time, facilitating training, automating record keeping, and increasing efficiency.  The video cameras constantly observe and record.  The way this system works is that it has been trained to recognize the steps in procedures undertaken by skilled technicians on the factory floor.  These technicians are taught a particular plan, composed of a set of steps, done in a particular order.  The system learns to recognize them.  If you think this sounds like it could lead to Taylorism run amok, you won't be the first.  Watching people and watching how they are doing what they are doing is important for safety and practical training but could also be seen as something that would provide management with an unwelcome gaze over the worker.

When we keep in mind the difference between a plan and a situated action (Suchman, 1987), we know that as good as a plan may be, a person may need to veer from that plan to account for local conditions.  So, when the local situation requires it, an intelligent being will find a way to reach the appropriate end state despite having to change some part of a plan.  This is not a situation that a typical deep learning-based system can account for.  One thing that a machine vision cannot do is to recognize something new.  It does not recognize novel actions for what they are, it simply recognizes that they are not the expected step.

Because of this, one of our roles here was to explain to management why they shouldn't always have access to what the machine "sees".  An example came up in our work.  During an observation, we saw an expert "going through the steps" when someone came up to them with a problem. This was standard protocol where someone with a problem should come to someone more senior for assistance.  This new problem was solved and the expert returned to his task.  This diversion would, of course, have caused the lengthening of the

time of that interrupted step, not to mention the overall process. Some members of management wanted to know what was happening every time the system didn't see what was expected but this is the sort of naïve error that would cause disruption in the work being done.

*Practically inscrutable*

Developers often say that one simply can't understand how a deep learning system works. It is difficult, to be sure, but the workings of the system could be understood. Jose Hanson (Hanson and Burr, 1990) argued years ago that because neural nets are implemented on state machines, we know that they *can* be understood: one state leads to the next by virtue of an explicit command and there is a set of input data; each can be clearly seen. It just takes a lot of time to analyze, a whole lot of time. It took Google weeks to figure out how AlphaGo came up with one of its moves and explain why it was able to beat the world champion Go player using that move. The important point, though, is that they *could* explain it. It *was* possible. It was just ridiculously hard. A non-expert could not be expected to interrogate a system in any kind of reasonable time. Experts can't even do this quickly. So, how do we interact with a machine?

With inexplicable ontologies derived from patterns in pixels, understanding is surrendered to a mostly "well-performing system" built in a way to ease machine processing.

## Human Scale: Description and Explication

Another way of looking at the previous examples is as a "failure of description". The system in the factory setting had an incomplete description of the technician's job. Going and helping another technician is actually a prescribed part of the job; it's merely infrequent. But as far as the system was concerned, prolonged absence from the process it knows is a problem like any other. So, a problem is seen where none actually exists because the system hadn't been trained to recognize this option (or myriad others). All of the possible actions that might be correctly undertaken by a technician are not possible to train the system to recognize because there are countless correct things to do and ways to do them. Instead, what the system can do is to learn a limited set of actions that could be undertaken and watch to see when they are done correctly. There are many valuable services that such a program can provide but 24/7 understanding of everything it sees is not one of those.

We also see this failure of description in the case of ASL. The system had not been trained to recognize the data relative to the categories of human perception, the types of data that make human language possible even when we're not explicitly aware of them. That is, phonemes and morphemes are important ways that humans see language even when we're not aware of them. And we noted that there are other types of data that the system won't see. Research in ethnomethodology, conversation analysis, and embodied interaction have demonstrated that we are signaling each other in many ways, often unconsciously, but those signals are nonetheless important for the interpretation of meaning. This may include such factors as subtle body positioning, direction, timing and coordination of gaze, and a host of other signals that happen too quickly or subtly to be easily described, but which nonetheless affect communication. The problem is, except for such micro-analytic work, those signals

are rarely even acknowledged and, to our knowledge, have never been included in a deep learning-based natural language system.

While (at least, heuristics for) each of these communicative categories could be learned by the system, it could only happen by resisting the emerging standard for such deep learning systems. Seeing ASL as a set of vectors of pixels, simply doesn't bode well for bringing this up to a human scale. Pixels are too fine a scale. Humans think of and see things in ways that are difficult to find in sets of pixels.

The challenge that failures of description present for deep learning systems, then, is that these systems will always be hamstrung.

A way out: changing how we design DL systems. Rather than designing a system as though it completely describes a process (e.g., servicing a tool or translating ASL), we should be developing systems that watch for events in the environment and provide further information in ways that recognize a potential insufficiency and are always compatible with the possibility of error. This is how we can provide a reasonable user experience in the face of deep learnings benefits and limitations. Spurious correlations will still happen, or maybe more correctly, "meaningless" event detection will still happen. Consider Geertz's discussion of the meaning of a wink (1973). Sometimes a wink will just be dust in someone's eye.

The implication here is that our DL systems must be bounded and targeted at the kinds of recognition tasks that make their level of activity more commensurate with human understanding and assessment. That means any individual DL system will perform a task that, if it produces a result that is not meaningful or useful to the human, the human user doesn't require hours of analysis to figure out what happened, but rather can disposition the result quickly, and in a way that the system can learn from.

## SUMMARY

The intent of this paper was to argue that one of the most significant recent directions in technology – deep learning – has flaws that are best addressed by those trained in ethnographic methods. Who better than ethnographers to advance the cause of human scale?

A generation (or two!) ago, ethnographers were brought into technology development in order to help people make products that fit people so that businesses could "scale up" their offerings and make them relevant for the whole world. However, once they were inside the corporation, so many more problems were revealed to be within the ethnographer's domain.

Atomic units are used to simplify programming. Pixels are used for images and spectrographic-style frequency analyses for speech sounds. It does simplify programming, too. It's just that it is the wrong level of abstraction for dealing with people.

The work we presented here was to say, in part, how we might create machine learning that works well but, beyond that, it's also about developing AI systems that can be more easily understood by people. Much of today's deep learning consists of the type of system that Latour could point to as being particularly rife with blackboxing (1999); because it is practically impossible to know how they work. Successful scaling-up of the technology should not mean that no one will have access to the methods behind the madness.

By getting the scale right for human understanding, we can hope to have more control over the gaze of the machine. This may slow down both processing and even system creation; it could even mean that a given system would not be as broadly applicable. But it

would be a better system, working at a more human scale, and would enable more fundamental interaction with the system itself.

**Richard Beckwith** is a Research Psychologist at Intel Corporation's Intel Labs.  He is a psychologist who studies the impact that emerging technologies have on those upon whom they emerge and helps to ensure that technology designs can support people in the way that they should.

**John Sherry** is the director of the User Experience Innovation Lab in Intel Labs. This organization focuses on the human dimension of machine learning technologies from diverse perspectives, to better imagine and prototype new technological possibilities, and anticipate the alignments necessary for those to become reality.


## REFERENCES CITED

Athalye, Anish, Logan Engstrom, Andrew Ilyas, and Kevin Kwok. 2018. "Synthesizing Robust Adversarial Examples". Accessed [24 Aug 2020]. https://arxiv.org/pdf/1707.07397.pdf.

Birdwhistell, Raymond. 1970. *Kinesics and Context: Essays on Body Motion Communication*. Philadelphia: University of Pennsylvania Press.

Chan, William, Nadeep Jaitly, Quoc Le and Oriol Vinyals. 2016. "Listen, attend and spell: A neural network for large vocabulary conversational speech recognition," IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Shanghai, 2016, pp. 4960-4964, doi: 10.1109/ICASSP.2016.7472621.

Console, Luca, Fabrizio Antonelli, Giulia Biamino, Francesca Carmagnola, Federica Cena, Elisa Chiabrando, Vincenzo Cuciti, Matteo Demichelis, Franco Fassio, Fabrizio Franceschi, Roberto Furnari, Cristina Gena, Marina Geymonat, Piercarlo Grimaldi, Pierluige Grillo, Silvia Likavec, Ilaria Lombardi, Dario Mana, Alessandro Marcengo, Michele Mioli, Mario Mirabelli, Monica Perrero, Claudia Picardi, Federica Protti, Amon Rapp, Rossana Simeoni, Daniele Theseider Dupré, Ilaria Torre, Andrea Toso, Fabio Torta, and Fabiana Vernero. 2013. Interacting with social networks of intelligent things and people in the world of gastronomy. ACM Trans. Interact. Intell. Syst. 3, 1, Article 4 (April 2013), 38 pages. DOI:https://doi.org/10.1145/2448116.2448120

Deng, Jia, Wei Dong, Richard Socher, Li-Jia Li, Kai Li and Fei-Fei Li.  2009. ImageNet: A Large-Scale Hierarchical Image Database. IEEE Computer Vision and Pattern Recognition, pp. 248-255, doi: 10.1109/CVPR.2009.5206848.

Dennett, Daniel. 1978. *Brainstorms*. Montgomery, VT: Bradford Books.

Feld, Steven. 1982. *Sound and Sentiment: Birds, Weeping, Poetics, and Song in Kaluli Expression*. Philadelphia: University of Pennsylvania Press.

Foucault, Michel. 1971. *The Order of Things: An Archaeology of the Human Sciences*. New York: Pantheon Books.

Geertz, Clifford. 1973. "Thick Description: Toward an Interpretive Theory of Culture." In *The Interpretation of Cultures*, 3–30. New York: Basic Books

Handle, Richard. 2009. "The Uses of Incommensurability in Anthropology." *New Literary History*, Vol. 40, No. 3, pp. 627-647.

Hanson, Stephen José & David Burr. 1990. "What Connectionist Models Learn: Learning and Representation in Connectionist Networks." *Behavioral and Brain Sciences*, 13(3), 471-489. doi:10.1017/S0140525X00079760

Harris, Mary. 2019. "Amazon Encourages Police to Use Untested Facial Recognition Technology." 24 May, 2019. Slate website. Accessed [18 August, 2020] https://slate.com/news-and-politics/2019/05/facial-recognition-police-officers-hillsboro-oregon-amazon.html

Kendon, Adam. 1967. "Some functions of gaze-direction in social interaction." Acta Psychologia 26: 22-63.

Krizhevsky, Alex, Ilya Sutskever, and Geoffrey E. Hinton. 2012. "Imagenet Classification with Deep Convolutional Neural Networks." Proceedings of the 25th International Conference on Neural Information Processing Systems - Volume 1, Pages 1097–1105.

Kuhn, Thomas S. 1962. *The Structure of Scientific Revolutions*. University of Chicago Press.

Latour, Bruno 1999. *Pandora's Hope: Essays on the Reality of Science Studies*. Cambridge, Massachusetts*:* Harvard University Press.

Le, Quoc V., Marc'Aurelio Ranzato, Rajat Monga, Matthieu Devin, Kai Chen, Greg S. Corrado, Jeff Dean, and Andrew Y. Ng. 2012. "Building high-level features using large scale unsupervised learning." In Proceedings of the 29th International Conference on International Conference on Machine Learning. Omnipress, Madison, WI, USA, 507–514.

Ludwig, David. 2016. "Overlapping Ontologies and Indigenous Knowledge: From Integration to Ontological Self-Determination." Studies in History and Philosophy of Science 59: 36-45.

Mahmood, Sharif, Sruti Bhagavatula, Lujo Bauer, and Michael K. Reiter. 2016. "Accessorize to a Crime: Real and Stealthy Attacks on State-of-the-Art Face Recognition." In *Proceedings of the 2016 ACM SIGSAC Conference on Computer and Communications.* 1528–1540. DOI:https://doi.org/10.1145/2976749.2978392

Malinowski, Bronislaw. 1984. *Argonauts of the Western Pacific: An Account of Native Enterprise and Adventure in the Archipelagoes of Melanesian New Guinea.* Prospect Heights, Ill.: Waveland Press.

Owens, Andrew and Alexei A Efros. 2018. "Audio-visual scene analysis with self-supervised multisensory features." Proceedings of the European Conference on Computer Vision. 639-658.

Schegloff, Emanuel. 1998. "Body torque." Social Research 65: 535-596.

Searle, John. 1983. *Intentionality: An Essay in the Philosophy of Mind.* Cambridge: Cambridge University Press.

Silver, David, Aja Huang, Christopher J. Maddison, Arthur Guez, Laurent Sifre, George van den Driessche, Julian Schrittwieser, Ioannis Antonoglou, Veda Panneershelvam, Marc Lanctot, Sander Dieleman, Dominik Grewe, John Nham, Nal Kalchbrenner, Ilya Sutskever, Timothy Lillicrap, Madeleine Leach, Koray Kavukcuoglu, Thore Graepel, and Demis Hassabis. 2016. Mastering the Game of Go with Deep Neural Networks and Tree Search. *Nature* 529, 484–489. https://doi.org/10.1038/nature16961

Suchman, Lucy. A. 1999. *Plans and Situated Actions: The Problem of Human-Machine Communication.* Cambridge: Cambridge University Press.

van der Wal, Reine. C., Robbie M. Sutton, Jens Lange, and João Braga, J. 2018. "Suspicious binds: Conspiracy thinking and tenuous perceptions of causal connections between co-occurring and spuriously correlated events." *European Journal of Social Psychology*, 48(7), 970–989. https://doi.org/10.1002/ejsp.2507

Vaswani, Ashish, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Łukasz Kaiser, and Illia Polosukhin. 2017. "Attention is All You Need." In *Proceedings of the 31st International Conference on Neural Information Processing Systems.* Red Hook, NY, USA: Curran Associates Inc. 6000–6010.

Watzlawick, Paul, Janet Beavin Bavelas & Don D. Jackson. 1967. *Pragmatics of Human Communication: A Study of Interactional Patterns, Pathologies, and Paradoxes.* New York, NY: Norton.

**THEMATIC SESSION**

# Built to Scale: Space is Different from Place

Architecture isn't just art; it physically shapes the ways in which we live and work. This session explores the intersections of vision and concrete building materials, thinking through the ways in which people and populations are affected by the built environment. In these ways, spaces are different from places. These works address the importance/centrality of different types of buildings (edifices) and the practices that they make possible.

Session Curators: Afra Chen, Lisa Kleinman, Carry Yury

# Ten People Thick

## Design for Change

SUE WITTENOOM, *The Soft Build*

*An exploration of scale in the built environment. Looking first at the graphic scale of building documentation - each layer with its own purpose and logic, being absolutely clear to reduce risk. And then considering architecture at human scale - how design thinking becomes a scaffold for organisational change. It explores how to engage with people in a visioning process, how existing environments shape world views, and how those conversations "scale up" from the individual, through the group, to form the aspirations for one new building, a multi-faith setting in Western Sydney that needs to find a new way to integrate Islamic and Christian theologies.*

*Keywords: User research, Organizational change*
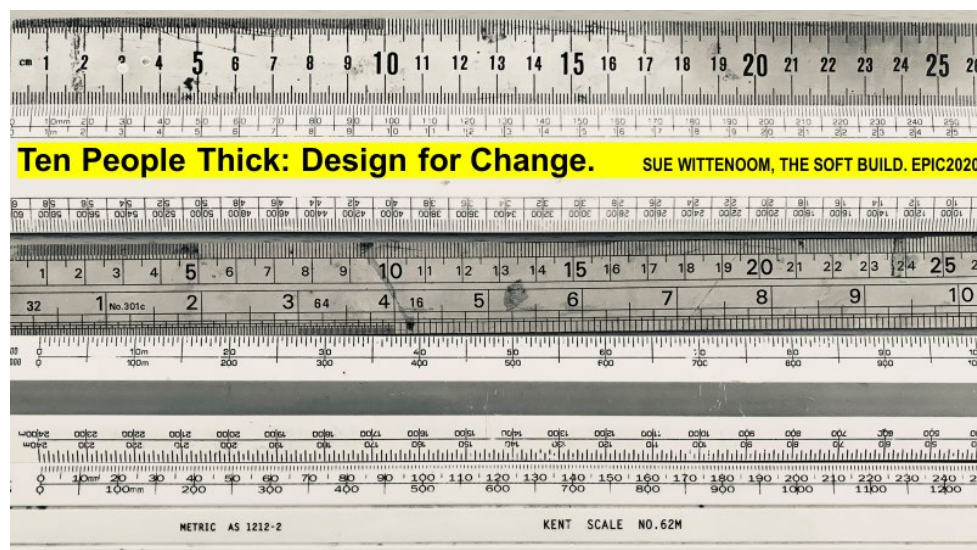


A lifetime collection of scale rulers © author

**Sue Wittenoom** is the founder of The Soft Build, a consultancy focused on design for change. A registered architect with an MBA,  her work has evolved over the past 30 years from architecture to project management, program design, change management and strategic consulting.  A director of global consultancy DEGW and AECOM's StrategyPlus practice in Australia until 2015, Sue now designs change frameworks for building projects that are new and different, using both the planning process and the built environment to shape new ways of living, working and learning. She is an experienced facilitator and design jury member, and a regular contributor to webinars, national conferences and industry forums on the impact of flexible working on our cities. swittenoom@thesoftbuild.com

CASE STUDY

# The Rollercoaster

## How to Go from Global to Local and Back Again—The Case of a Walking Drive Model in Paris

MARC-ANTOINE MORIER, _unknowns

*The act of shopping for food is a very local experience, yet large food retail chains have built their business on homogenizing and standardizing the experience. In this article, we mobilize an ethnographic study carried out in 2018 for a food distributor regarding a new model of online retail pick-up. The goal of the project was to understand how a new method for food shopping could be scaled across different types of neighbourhoods. We created a scale model that incorporates both individual shopping practices and the demographics of the neighbourhood; using ethnographic methods as the basic unit. Using concepts from gentrification, we also contextualize our insights within the changing dynamics of a neighbourhood—because places are not static entities. We discuss how the scale model could be used to duplicate results from one neighbourhood to another and the reception of our work by the client.*

*Keywords: scale, retail, walking drive, gentrification*

## THE BUSINESS ISSUE: FROM HYPERMARKETS TO LOCAL "WALKING DRIVES"

### How to Position and Present Ethnography in a World That Doesn't Know It

February 2019, Paris, France. The Director of Innovation of the world's second largest food retail group (€2,1 million sales annually, with 12,300 stores in 30 countries) questioned us about the value of the online retail pick-up points they'd started experimenting with in urban areas. The walking-drive model is simple: a counter, not located in a hypermarket, where online orders can be retrieved by customers. Their idea was simple: to improve this model, we needed to start from the needs and expectations of urban customers in terms of food shopping. And since not all urban customers have the same needs, the "walking-drives" should meet local expectations.

Historically and culturally, however, nothing predisposes this group to reason in this way. As a central player in the food industry since the 1950s, the company is based on this triptych: *mass* production, thanks to a food industry with productivist breeding...*for masses,* i.e. international "markets" where large-meshed typologies are addressed ("young", "old", "rich", "poor")...*with massive means* (giant infrastructures, organizational standardization, macro indicators—and the eternal myth of food abundance).

This logic culminated in the invention of hypermarkets in the 1960s. A symbol of mass distribution in France, this model has since been the subject of much criticism. More precisely, this all-scalable logic is regularly accused of "killing the local". By reproducing "miniature towns" inside their shopping galleries—which are set up at the entrance to

towns—hypermarkets are said to contribute to the closure of small shopkeepers (grocers, butchers, greengrocers, hairdressers, etc.). With online shopping, they would also capture part of the social life of the town centres that is based on commercial exchange. Shared and relayed by many local elected officials, this fear gradually reached the scale of a public order problem, to the point of triggering intervention by the French State in the most affected localities[1].

More recent but also more confidential, another criticism has been made by French anthropologist Marc Augé (1992) with his concept of non-places. A non-place is a space with no history, no identity, and no social relations. In other words, it is an interchangeable space in which the individuals who use it remain anonymous bystanders. According to the author, these spaces proliferate, while "folklorizing" local identities: they include airports, train stations, department stores and, of course, hypermarkets. Shopping galleries are a good example. Hypermarkets have recreated the markets of the city centres by removing its disorder (the auction, price negotiation, smells, unruly crowds, winding alleys, etc.), by imposing standards (storefronts must all be the same size for example). They however kept the city centre's promise of localism, "authenticity" and proximity. But, in the end, every shopping gallery is the same, wherever they are.

To sum up, on the one hand, hypermarkets are accused of siphoning off that part of the social life of city centres that is based on commerce. On the other hand, they are suspected of weakening the part of socialization that is based on space. Wherever they settle, they contribute to the desertification of city centres, by proposing as an alternative trade without the social relations and local histories attached to it. With their scale logic, they install clones in the strongholds of particularism. Scaling, as Anna Lowenhaupt Tsing (2015) writes, would therefore be to "eliminate diversity"?

Here's the set for our ethnographic work. So the position our client suggested seemed relatively new because he took these criticisms seriously. Nevertheless, his motivations were more prosaic: at the beginning of 2019, after one year of launch, the results of the walking-drive (subsequently referred to as the Drive) were mixed. Although the Drive had attracted customers, attendance and average basket size were perceived as insufficient. Numbers didn't live up to expectations. To solve this issue, the Director of Innovation took the way the group had built itself in reverse: starting from the local characteristics of a place to propose an alternative food range or to build new non-food service offers; and not building service offers and then putting them in a place, whatever its characteristics.

Following a call for proposal, the food retailer's innovation team selected _unknowns because the methodology we created allowed us to study a specific neighbourhood and made it possible to scale the results produced to other neighbourhoods. Our client wanted us to create 2 monographs. The first one was to be done in the Parmentier in the 11th district located in the east of Paris because this is where the new Drive was located. The second one was in a perimeter around a store located in Villeurbanne, a city in the middle of eastern France, connected to Lyon, because a second Drive could be deployed there. We mainly describe the survey conducted in Paris.

## APPROACH

### The Research Issue: Typical Cases x Similar Urban Characteristics = Scalability?

With socio-economist Max Weber's teachings (1986), we have become accustomed to characterizing the city in a very different way from what scalability implies. What is striking in European cities, and arguably elsewhere, is their cultural, political, legal, and economic specificities. The same is true at a lower level—what characterizes neighbourhoods is a topography (a hill, a river), a dominant function (recreational, residential, intellectual, economic), an architectural style, a historic event (La Bastille in Paris), an emblematic personality (Authier, 2006). And above all, a subgroup of the general population (student, executive, couple, retired, etc.) with specific needs.

We had to understand those specificities to analyse their impact on the "food races". In order to leave the high & macro scales and go down to "human height", the ethnographic approach, by the concern it brings to the description of details, was our best ally. However, we also needed to reintegrate our teachings into a scalability scheme. The 1:1 scale of the ethnographic study suddenly seemed too narrow. In other words, we had to find a way to take the service offer imagined on the basis of a neighbourhood and duplicate it in another neighbourhood; potentially in a different city. But how could we scale a walking-drive built on hyper-local singularities? How could we scale the "non scalable"? We had to find a unit of analysis with a better potential for generalization. Figure 1 shows the generalized framework we built and our results.
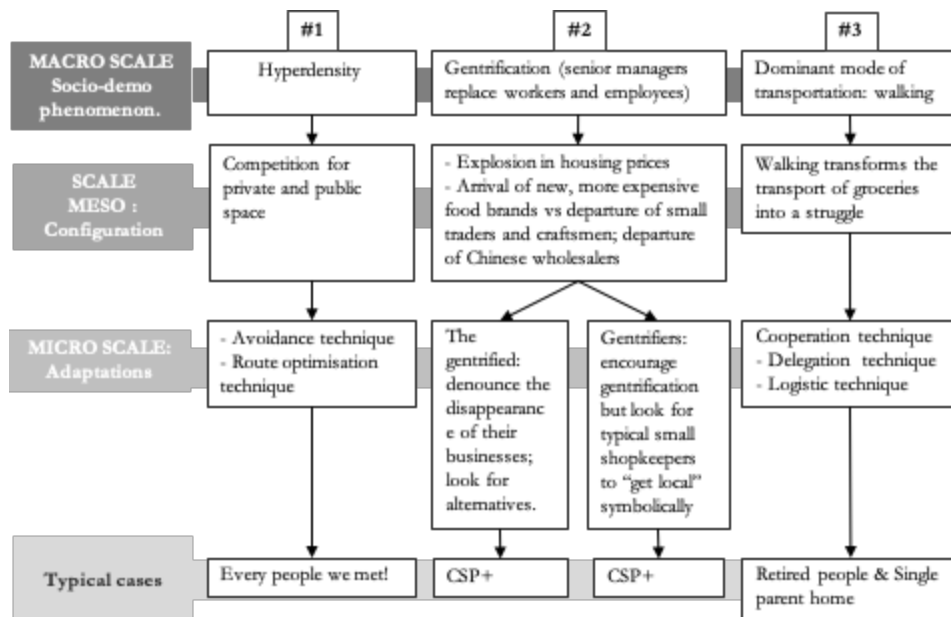


Figure 1. Generalized Scale Framework ©_unknowns

Let's explain each level.

*The Macro Level*

To begin with, we completely changed our focus: like an entomologist looking at an insect with a magnifying glass, we went up a notch to get a view from above, while diving into the details.

This leads to questions such as:

- Who are the inhabitants?
- What are their professions?
- What degree do they hold?
- How much do they earn on average?
- etc.

To get this macro point of view this, we identified and then examined statistical and cartographic studies from the national census, data from the Ateliers Parisiens d'urbanisme (APUR) and data on the evolution of prices per m² from the Notaires de France. This enabled us to discover, it was then a presupposition, that within the 11th arrondissement, the population did not have homogeneous characteristics, and that consequently, subgroups were distributed differently in space. For example, on the map in Figure 2, published in 2012 by the National Institute of Statistics and Economic Study (INSEE), we see that middle class people (in orange) live side by side with executives (in yellow) along with people having several social backgrounds (in green; i.e. students, young graduates, migrant workers, recipients dependent on social benefits, intellectuals, etc.).
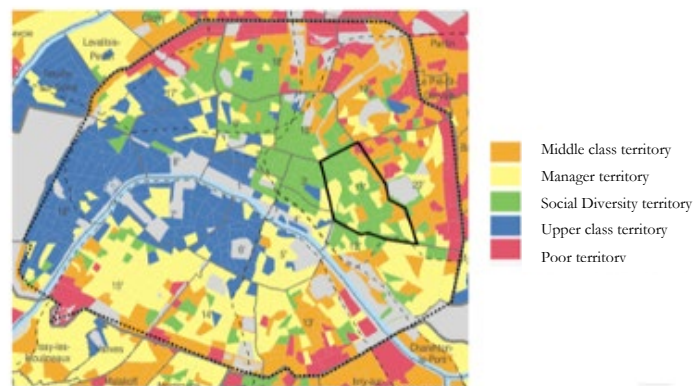


Figure 2. Neighbourhood cartography. Adapted from ©INSEE

However, we lacked a framework for interpreting these data. Urban sociology offered us the concept of "gentrification" which we will see later. Intuitively, we knew that this concept made it possible to read the city not only as pure physical data, flows in a topography, but as the "projection on the ground of social relationships" as French geographer Henri Lefebvre

once said (1968). In short, the neighbourhood we were to study could be read as an arena where two sub-groups—the gentrified and the gentrifiers—were struggling to appropriate a neighbourhood; theirs in this case. One of our hypotheses was that food stores played a role in this struggle.

*The Meso Level*

Like town planners who wonder how people move around according to the transport infrastructure, we asked ourselves how people shop according to the urban characteristics of the neighbourhood where they live.

This leads to questions such as:

- Do the buildings have stairwells wide enough to carry several bags of groceries?
- Are there any level breaks on the roadway?
- How easy is it to drag a shopping cart down a crowded street?

There was also a need to find a way to observe eating routines, such as lunch breaks for employees, and to track their deployment in space—for example, the use of a park. The objective was therefore to uncover the encounter between urban space and shopping practices. This is what we called configurations, or the meso level.

But how do we capture it? In the beginning, the temptation was great to want to "observe everything". But this dream of ubiquity, already illusory in the context of a restricted observation perimeter (a schoolyard for example), became impossible on the scale of a neighbourhood. In the manner of ornithologists who want to observe the passage of migratory birds, we did some spotting to identify observation posts. School outings, parks, subway exits, and pedestrian walkways seemed to be the most promising places in terms of feeding routines. In order to increase the hourly scope of the observations, we also took accommodation on site for the duration of the study.

*The Micro Level*

Thus, in order to combine the macro and the meso scale with the 1:1 scale of the field, we decided to reuse the concept of gentrification to recruit respondents for the study. In other words, we broke down the concept of gentrification into socio-demographic characteristics in order to recruit people according to whether they were gentrified or gentrifying. Since the type of profession (occupation groupings from *Catégories Socio-Professionnelles* CSP level 1) is the criterion that is most likely to differentiate one from the other, we made it a central recruitment criterion. Thus we decided to recruit by "ideal-typical situations", in order to catch general social processes (Becker, 2014; Passeron, 2015), in particular gentrification (Clerval, 2013). For instance, we assume that an artisan, who does not own his or her main dwelling, and has a low income tax rate is typically a gentrified person. Conversely, we would have a chance to find a gentrifier by recruiting an executive who owns his own home and whose tax bill is high. The micro level raises questions described in Table 1.

**Table 1. Questions at the Micro Level**

| For the gentrified | For the gentrifiers |
|---|---|
| • How do you shop in a neighbourhood where prices are increasingly growing?<br><br>• Where do you go? Do you stay in Paris? For which products?<br><br>• How do you get them home? | • Did you pay attention to the businesses in your neighbourhood when you moved in?<br><br>• Are there stores that you never visit?<br><br>• Are there shops that you would like to see more of? |

In the end, our "scaffolding" brought together different scales of analysis, the combination of which should make it possible to fill in the blind spots specific to each one:

- The macro scale counterbalances the empirical ground level by allowing us to detect social processes nested in an observation or an interview—which are invisible to the naked eye.
- Then, the meso scale makes it possible to take into account the influence of urban characteristics on shopping for food practices; a spatial dimension that is difficult to capture by statistics alone.
- Finally, the micro scale captures how individuals experience these general social processes, which are invisible from the top of the scaffolding.

Our hypothesis for scale was then to compare two monographs: If the characteristics of neighbourhood A (e.g. gentrification) could be observed in a neighbourhood B, our teachings, fueled by observations, interviews, and other statistical and cartographic data, would be valid. Therefore, in our model, a teaching is "valid" when it is observed in two different monographs. In other words, where the results were identical, we could duplicate the new offers; for example, a concierge service or a new food range. Where they were different, we couldn't do that, or, at least we would need to restart a study. The equation is as follows: if *typical cases x urban characteristics of A = typical cases x urban characteristics of B* then the model is scalable.

## RESULTS

### Resources Under Pressure

Looking at INSEE's statistical data, one of the first things that struck us was the extreme density of this neighbourhood. In 2010, 44,744 inhabitants lived there per km². In other words, there are twice as many inhabitants in this neighbourhood as in the rest of Paris on average (21,200). By way of comparison, there are 7,100 inhabitants in New York City per km² on average in the same year.

But after all, why is this a problem? It's a problem because it means that demographic pressure is putting pressure on the resources located in this territory—the foreground space. Hence, for example, policies to de-densify the territory, as shown in Figure 3 with what urban planners call a "green tooth", i.e. public gardens installed between two buildings.



Figure 3. Installation of a public garden on Voltaire Boulevard,
in front of Saint-Ambroise Church. ©_unknowns

Said in less policed terms, it means that residents are competing for the space and facilities there: housing, green spaces, parking, and of course food stores. For example, one need only look at the neighbourhood's public library to observe traces of saturation; such as this calendar posted at the entrance to a public library that informs about usage levels (see Figure 4). By walking around the shelves of this same library, one understands something else: to avoid being deprived of available resources, some residents bypass the commonly established rules for sharing these resources. They do this in order to capture resources before they are captured by others—in this case cultural goods such as DVDs as shown in Figure 5.
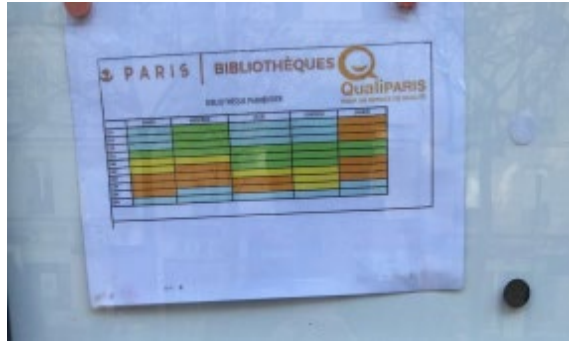
Figure 4. Poster representing the peak hours of the Parmentier media library. ©_unknowns



Figure 5. Media library of Parmentier, on level -1, in the poetry corner, a little isolated from the main aisles. The sign tells patrons not to hide the DVDs "behind the books." ©_unknowns

Another characteristic of the neighbourhood is the relative difficulty of getting around. Coupled with the high density, the narrowness of the sidewalks as well as the level breaks make the transport of shopping a real ordeal. So much so, that when faced with a commodity, some inhabitants estimate the effort required to bring their goods home before deciding that when on foot, they will be selective about what they purchase. In the end, the answer to this question has of course an influence on the choice of store. But it also gives rise to tactics to reduce the drudgery—such as the interviewee who buys heavy goods only in the grocery store downstairs. Others divide up the carrying work, such as the interviewee who asks her neighbour to go to the store with her to help carry water bottles home. From this perspective, helping individuals with their errands means reducing a constraint produced by the meeting of demographics (high density) and urban characteristics (narrow sidewalk, level breaks).

**Thinking Customer Segmentation Through Gentrification.**

Up to now, we have talked about competition from residents without really specifying the identity of the protagonists. Who are these inhabitants? How do they form sub-groups? And above all, do they have different needs in terms of food shopping? In order to find the most differentiating marker possible, we used the concept of gentrification, or *embourgeoisement* in French. This notion comes from the Anglo-Saxon geography of the 1960s.

To my knowledge, it was the sociologist Ruth Glass (1962) who first used gentrification to describe, in the neighbourhoods of Notting Hill and Islington, the transition from a working-class population to a more affluent population, the gentry.

The lens of gentrification allows us to sociologize our analysis a little more. It can now be hypothesized that competition for resources is a social competition, bringing together social groups that do not have the same characteristics and therefore may not always have the same interests.

Statistics about changing professional classes also established that gentrification occurred in the neighbourhood. If we look at them, we learn that between 1954 and 2010, the share of Executives and Senior Intellectual Professions, Business Leaders, and Intermediate Professions increased by 45 points, from 28.6% to 73.8%. This is exactly the number of points lost by the share of blue-collar and white-collar workers over the same period: from 71.4% in 1954, to 2010 representing 26.2% of the population of the 11th district. This inversion continues today. If we compare only workers and executives and higher intellectual professions, we can see that between 2010 and 2015 the share of the former is decreasing (from 5.1% to 4.3%), while the share of the latter is increasing (from 30.3% to 32.5%). In short, managers are the majority in the district and blue-collar workers are the minority professional class.

If we take it down a notch further, at CSP level 2 which are trades professions, this means that garage owners, masons, craftsmen, cobblers, upholsterers, printers, and metalworkers have gradually given way to artists, production managers, association leaders, theatre company administrators, nurses and secondary school teachers. This was what we could call the first wave of gentrification. In a second phase, senior executives arrived in the neighbourhood: they were more likely to be professionals (lawyers, doctors, company directors) or private sector executives (consultants, senior managers, etc.). In 2015, higher education graduates represent 2/3 of the population (61.5%). These sociological changes are modifying the supply of catering and food consumption in the district. Figure 6 shows an organic grocery store on avenue Parmentier which replaces a low-cost Franprix market.



Figure 6. Bio c'Bon organic grocery store near the Drive on avenue Parmentier. ©_unknowns

If we zoom in a little more, this time at the individual level, we can see that these changes are assessed differently. On the one hand, gentrified people castigate these changes because they see small traders disappearing in favour of restaurants. Figure 7 exemplifies a

gentrified space that's opened in the neighbourhood. A resident in Voltaire who is a receptionist at the Maison des Associations describes this change:

> "We used to have a new food trader [greengrocer, butcher, fishmonger] every week—[She turns around and shows me the shops in front of Maurice Gardette Square]—now we have no more shops. Now it's just restaurants, look: 1, 2, 3, 4, 5. We've got more than that. And rue Saint-Maur is just that. Only bars and restaurants."



Figure 7. Cocktail bar and open space La Popina at rue Saint-Maur. Inside, a white man of about 40 years old in a white V-neck t-shirt consults his iPhone. A Mac decorated with a sticker is placed in front of him. ©_unknowns

Clothing wholesalers crystallize the opposition. For the poorest gentrified, they are an opportunity to buy affordable clothes—that is to say, to control their spending. As a saleswoman in a jewellery shop explains, "It's tempting because there are some interesting items, eh? I've tried, but no, no, no, we don't buy retail. But they have some nice stuff." On the other hand, for the gentrifiers, these wholesalers have to close down to make way for shops more in line with their taste, that is to say with their social position. As the director of a business school in the 12th arrondissement explains:

> "On Boulevard Voltaire, all the Chinese wholesalers are leaving. I hope I don't have to tell you this, but they are being replaced by shops. We are very curious to know who is moving in. [...] In fact the Marais, finally the transformation of Beaumarchais must come to Voltaire. [What shops do you like in Beaumarchais?] It's clothes shops, it's APC, the Blend restaurant [...] all the brands we like Maje, Bonpoint for children."

This was the first interview of the study and it seemed to us emblematic of the more global process of modification of the sociological composition of the neighbourhood. In the end, it could have been called "extraordinary gentrification calculation" because our business school-educated director finally had a winning speech: not only did he want to take advantage of the effects of gentrification, but also to multiply them.

Obviously, the installation of these new populations is not without opposition. By settling, the newcomers also install new rules: what can be allowed in the neighbourhood or what is no longer possible. These new rules are sometimes denounced by the former inhabitants, who feel like they are "dispossessed" of their former stronghold. An association

leader described her outrage that the *bourgeois* call the police when young people play football outside the hours set by the town hall: "[And the population, you've seen it change in recent years...] But yes, even in the square, Maurice Gardette, there are obnoxious people. They are the bourgeois who want order."

But what do food races have to do with it? In fact, the establishment of this type of business may give the gentrifiers hope of attracting their fellow people, i.e. other executives, other engineers, or other lawyers. And thus, strengthen their presence in the neighbourhood by multiplying the small bastions in order to occupy the space. Here they will be able to live out their social status.

In the long term, it is a question of increasing the added value of their residence when they sell it. In other words, from this point of view, gentrifiers have every interest in ousting businesses that do not inspire confidence among future buyers who look at the type of store in a neighbourhood to decide whether or not to invest there. This is an indication of the progress of the gentrification front. From this point of view, an organic store is a favourable index; a discount store is an unfavourable index.

But this eviction should not be total. The geographer Anne Clerval (2013) points out that newcomers to the working-class districts of Paris also need to stage their anchoring in the neighbourhood they have newly moved into.

> "The frequenting of small shops gives the gentrifiers the impression that they are participating in the sociability of the neighbourhood (164) [...]" despite the social and cultural differences that they import there. In other words, it is a means of capturing symbolic profits, those offered by the reputation of being "open-minded."

On the nice side, however, they fear that they can no longer afford to live in their neighbourhood because the price of housing and various goods is rising faster than their wages. Staying in the 11th arrondissement therefore forces them to invent different schemes. As far as food shopping is concerned, we met gentrified people who simply stopped shopping in the neighbourhood. They now have to get their supplies elsewhere in Paris, i.e. where gentrification has not yet arrived:

> "The last purchases I made on special offer was dishwashing liquid; in normal times it's between 1.60€ and 1.80€ and on offer it's 3 for 3€—so if there are bargains in Auchan [in Bagnolet], often it's on Wednesdays, I go there…I also look on the Internet every Sunday, I look at all the signs. That's how I do my shopping."

If we think of food shopping sessions as "acts", then we can say that the gentrified are gradually becoming deprived of a means of asserting their belonging to their neighbourhood every time they give up.

## DESIGNING FROM SCALE: NEW SERVICES AND ADAPTED POSTURES

### From A Business Point Of View

The adaptations observed in the typical cases were problems to be solved in people's lives—the basis of the future Drive offer. It was seen that some properties were difficult to access; because of the number of people who wanted them. What's the big deal? What can a food distributor do with this information? In fact, it means that one of the possible ways to expand the Drive's offer would be to offer the goods that are locally the most under pressure. It could be workspace, the very one that's taken over at the library; or, to use our example of cultural goods, books and DVDs. Thus, while generating additional traffic for the Drive, it would make the resources that are most in demand at the neighbourhood level a little less scarce. To paraphrase Hobbes, it is a competition of all against all that the new offer could help to "relax" and thus be a solution to the problem of hyperdensity.

We also saw that the transport of groceries was difficult. Another solution could have been to offer home delivery services for the heaviest and/or most bulky goods. Another would have been to lend shopping carts, cargo bikes, or even an electric scooter, in exchange, for example, for a subscription to a loyalty card and/or a deposit. Here, the Drive allows to delegate the carrying of shopping or to equip the customers to reduce the drudgery of this task.

Concerning the problem of gentrification, several options were possible. On the gentrified side, we saw for example that, because of the closure of their shops, the poorest people had to go to the outskirts to find affordable prices. To enable them to stay in their neighbourhood, the Drive could have offered a range of "essential" products sold at lower prices than in the new shops. Similarly, services could have been devised to increase their income; for example, by mobilizing their assets. These could have been services to facilitate the seasonal rental of their property, be it the handing over of the keys to the tenant, a cleaning service, home improvement to be carried out, etc.

On the gentrification side, it could have been to move upmarket, with organic products, for example—capable of distinguishing them socially, while affirming their belonging to the neighbourhood. Another manifestation of gentrification could be to install the iconic markers of the neighbourhood's history (a zinc counter, for example, typical of the former workers' bars in the neighbourhood). In addition, services facilitating their identification with the neighbourhood could have been imagined; for example, by positioning themselves as trusted third parties between them and different trades, especially craftsmen. A mixed offer could also have been imagined, aimed at both gentrified and gentrifying people, so as to position itself as a place of "cohabitation," meeting the needs of both groups would be an eating area on the small square facing the Drive.

### From a Design Point of View

Taking those local specificities into account could avoid building "non-places" devoid of any identity, such as hypermarkets, because of the scalability model they are based on that erases local specificities. For example, we could imagine designing a store integrated into the

local history of the neighbourhood, or an offer adapted to its main functionality (for example, commuting to Parmentier).

As to the impact on the client itself, we tried to shift

- ***posture***; no longer thinking of its walking-drives only as the end of a supply chain, but as a neighbourhood business, with social issues at stake,
- ***reading grid***; rethinking its customer segmentation in terms of gentrifiers & gentrified, in Paris, and adapting its range according to their needs,
- ***design method***: offers of services thought from a *typical cases x urban characteristics* grammar), and finally,
- ***deployment method***: if *typical cases x urban characteristics of district A = those of district B* then we can deploy.

## AREAS OF IMPROVEMENT

In hindsight, I would have done three things differently.

### Study Protocol and Analysis

First, not all observation units were comparable. French national statistics take the household as a unit, whereas the interviews conducted had each member as the unit. The former obliterated the distribution of domestic work within the couple (especially shopping) but not the latter.

Second, although the articulation of macro-meso-micro scales makes it possible to make otherwise scattered facts intelligible, it sometimes resembles a cosmogony that is too coherent to be true (De Sardan, 1996)! This raises the question of the place of scale comparison in the research process: is it the end or the beginning of research?

In other words, I think the scaffolding has been effective in making assumptions. It was "heuristic" as sociologists say. But to think that all the scales easily fit together, that you only have to pull the ball of wool to see all the dimensions of a subject—in this case demographic, urban, social, and finally political—is to give too much credit to the idea that the phenomena we observe are all consistent with each other, whatever the scale on which they occur.  In any case, the interweaving of these different levels of analysis is as much a datum to be explained as it is an explanatory datum.

Another disadvantage linked this time to the problematization in terms of gentrification. If the advantage is not to be naïve about the social issues behind the census figures, one pitfall is to take the side of the gentrified or the gentrifiers without realizing it. We are tempted to be miserable when we talk about the gentrified, we are tempted to be accusatory when we talk about the gentrifiers. In other words, without being careful, we can pass imperceptibly from a judgment of fact to a judgment of value.

### The Implementation of the Scale Model

The scalability model we have proposed (*typical cases x urban characteristics*), forces us to work with clones, i.e. similarities in similar contexts. This puts considerable weight on

variables from the first monographs (e.g. *hyperdensity + gentrification + walking*). That said, these variables can be verified fairly quickly—provided, however, that there are agencies that produce these statistics. In countries where such agencies do not exist, it may not be possible to collect this information.

## Receipt of Results at the Client's Premises

We have presented our results to several directorates. However, the higher we went up the hierarchical levels, the more the macro reference scales resurfaced, with the representativeness of the study as a banner. One of our mistakes was to think that the ambition of the Innovation department was shared by all stakeholders. And that everyone around the table was ready to integrate into their usual client typology, a typology inspired by the Marxist geography of the 1980s. This is probably forgetting that our interlocutors were not politically neutral. Even taking all possible precautions to de-politicize this notion, I am not sure that our client has appropriated it as a tool for description and analysis. In other words, an analysis in terms of gentrification was probably too "radical" to be accepted—at least we should have formulated our ideas differently.

Another issue was the unit of analysis. As a unit of analysis, the neighbourhood seemed too small compared to the units of analysis that were usually used; typically a city or even the region. However, it was the highest macro level of our framework. Even when proposing a scalability model to deploy the new offers, this spin-off logic was too different from the usual model of scale they use; i.e. to duplicate the same offer everywhere with a few adaptations at the margin according to local specificities.

In the end, perhaps everything was too new in this study and we should have said from the very start: "This is the first time they're going to hear about social science, gentrification, a swarming scale model, etc." We should have better adapted our discourse to our interlocutors. To do this, we should have consulted more with our innovation interlocutor.

**Marc-Antoine Morier** is an Anthropologist & Strategist at _unknowns, innovation consultancy *marcantoine.morier@gmail.com*

**NOTES**

1. During a visit to Girancourt (in eastern France) in 2010, former Prime Minister Edouard Philippe (Macron 2017-20 government) described the closure of cafés and bakeries as a "silent and banal

disaster". Tax exemption measures for local shops were therefore announced. Cf : https://www.franceculture.fr/emissions/les-bonnes-choses/les-magasins-ont-ils-un-avenir.

## REFERENCES CITED

Augé, Marc. 1992. Non-Places: An Introduction to Anthropology of Supermodernity. Le Seuil

Authier, Jean-Yves and Bacqué Marie-Hélène and Guérin-Pace France, Le quartier. Enjeux scientifiques, actions politiques et pratiques sociales. La Découverte. 2006

Becker, Howard S. 2014. What About Mozart? What About Murder? Reasoning from Cases. The University of Chicago Press

Clerval, Anne. 2013. Paris sans le peuple. La gentrification de la capitale. La Découverte.

De Sardan, Olivier. 1996. Abuse of data. Concerning some forms of overinterpretation in anthropology. Enquête, 3

Glass, Ruth. 1962. London: Aspects of Change. MacGibbon & Gee

Dumez, Hervé. 2015. "What is a case, and what is a case study?". Bulletin de méthodologie sociologique/Bulletin of Sociological Methodology. n° 127. pp. 43-57

Hobbes, Thomas. 1985. Leviathan. Penguin Classic

Lefebvre, Henri. 1968. Le droit à la ville. Éditions Anthropos

Passeron, Jean-Claude. 2015, Pensez par cas. Éditions de l'EHESS

Lowenhaupt Tsing, Anna. 2015. "Some Problems with Scale". The Mushroom at the End of the World: On the Possibility of Life in Capitalist Ruins. Princeton University Press

Weber, Max. 1986. The City. Glencoe

CASE STUDY

# Who Cares Where?

## A Pivotal Ethnographic Study for Italian Hospital Homecare

ISABEL FARINA, *Experientia srl*
ELENA MESSINA, *Experientia srl*
MICHELE VISCIOLA, *Experientia srl*
ELENA GUIDORZI, *Experientia srl*
CHIARA AGAMENNONE, *Experientia srl*
DARIA CANTU', *Experientia srl*

*The case study presented is an in-depth view on the project "Casa nel Parco" (translated as "the House in the Park"), a three-year, European-funded project (ERDF Funds 2014-2020) in the Italian region of Piedmont that involves 4 hospitals, 2 large companies, 14 small-medium enterprises, 2 universities, and 2 private research centers. The goal is to research and innovate hospital-homecare services for elderly and ALS patients, as well as their caregivers, through the implementation of e-health solutions. The uniqueness of our case study lays on the fact that our ethnographic work was pivotal in shifting the narrative of closed hospital ecosystems (Goffman 1961); where those outside of the hospital environment are not viewed as credible or essential sources for improving the care system. In this study we share how we built trust and negotiated a complex network of stakeholders and technical systems to successfully influence the design of homecare services. Ethnography played a major role in identifying user requirements (patients, caregivers, healthcare professionals) and in building scenarios for product and service prototype development, technology adaptation, and testing. In this way, ethnography and design were used as holistic and critical approaches to addressing health challenges and change. Successful implementation could therefore be considered as a socio-technical design challenge, rather than a pure technological design challenge, emphasizing a non-conflationary approach in which the social and the material are held apart for the purpose of exploring their interplay (Mutch, 2013). The key success that ethnography brought was in identifying how technical systems that were designed in the abstract could be used for the very real and specific problems faced by the elderly and ALS patients needing homecare.*

*Keywords: hospitals, ethnography for technology, EU-funded project*

## WHY IS IT IMPORTANT TO SPEAK ABOUT HOSPITALS AND HOME CARE IN ITALY?

To begin, it is important to recognize the lack of existing literature regarding ethnographic work within funded hospital projects in Italy. The literature regarding ethnography in hospitals outside Italy (Goffman, 1961; King, 1962; Van der Geest & Finkler, 2004), identifies two main aspects, which are also relevant for the Italian context: 1) reluctance by the hospital to open space for observation and ethnographic research and the fact that 2) hospitals are closed institutions with no fluidity between the inside and the outside or among the stakeholders dealing with its particular setting. Using ethnography to

identify user requirements is *per se* innovative considering the Italian cultural context. The ethnography conducted focused on two different hospital departments that provide homecare services in the Piedmont Region: the geriatric department for acute care at the Molinette Hospital in Torino and the Regional Expert Center for ALS at the Maggiore Hospital in Novara (CRESLA). The patient services offered by these hospitals are unique in that care also takes place outside the hospital walls. They reflect a willingness to shift the model of care, made by the hierarchical doctor-patient relationships, by integrating important actors within the model, such as caregivers and the home setting (Milligan, Roberts, & Mort, 2011). However, despite the excellent results, this service still encounters major resistance in terms of transferability to other clinical settings. Indeed, literature shows that too often, older and more fragile patients are excluded from telecare system design and implementation is often wrongly seen as a one-off event (Delle Fave & Marsicano, 2004). In contrast, the project presented here aims at supporting safety, autonomy, and the independent living and welfare of frail, old and disabled people (Orlikowski & Robey, 2001). The current viral epidemic in Italy, which has hit the country's northern regions in particular, has demonstrated the need for good, alternative models of care that are able to change the health ecosystem.

## HOW TO POSITION AND PRESENT ETHNOGRAPHY IN A WORLD THAT DOESN'T KNOW IT

Clinical settings and the field of medical device engineering are not comfortable dealing with qualitative research, due to their resistance in allowing outsiders to observe internal modes of working (Zaman, 2004). In order to perform ethnography in a medical setting and deal with diverse stakeholders who do not necessarily know what ethnography is, we had to build trust. We also had to educate people about our research methods, data analysis, and goals. This was primarily done by using a participatory and collaborative approach able to respond to any doubts, perplexities, and needs of the partners and stakeholders.

Diverse stakeholders may hold different assumptions, values, and worldviews. They may also 'talk past' one another and compete for recognition and resources. In order to achieve the project goals, a more effective inter-stakeholder dialogue was set up to establish an organizational vision that better accommodated competing discourses. Performing stakeholder interviews was the best way to engage with project partners—as said, there were many and not all of them were clear about our tasks for the project. The word *stakeholder* is here intended to refer to the 25 technology partners that were part of the project and the two hospital centers considered for the research. Secondly, interviews allowed us to understand the initial assumptions regarding users. This was very important because the different technologies considered for the project were born originally for different purposes and users. For example, we understood that most of the stakeholders were not clear with the idea that patients under the care of the geriatric department were in an acute state, meaning that they were heavily sick, most probably with Alzheimer's in a serious stage. They were expecting active elderlies, able to self-manage their care (Milligan, Roberts, & Mort, 2011).

Hospitals and centers were both stakeholders and final users of the technologies in place. We performed stakeholder interviews as exploratory research to set expectations and to establish rapport. After that we started working closely with doctors and nurses in the design of the research protocol. This collaboration changed our initial idea of ethnographic

research. From a methodological point of view, we had planned an ethnography based on shadowing in both hospital and home environments, interviewing selected caregivers and patients according to a rigid scheme of inclusion criteria. With shadowing, our idea was to spend some time at the hospital to observe nurses' work, a typical workday, and documenting patients' households with pictures. We had to redefine our methodology in order to be more flexible on recruitment and to find a way to observe nurses' work and spend time with caregivers. We also had to conduct interviews with the counselor of the hospital homecare service as a representative of the hospital.

We also experienced more reluctance from one hospital department, which was supposed to be the third field where we planned to conduct our research. But unfortunately, we haven't been able to present our methodology and negotiate ways to overcome their fear of letting us in. So, we never performed research in that department. Finally, we had the chance to interview 5 caregivers and 2 patients, and 5 healthcare professionals from the OAD service and 6 families and 10 people of the clinical staff from CRESLA.

The opportunity to create a bridge between the hospital staff and our team has allowed us to enter the field more easily and perform research while accompanied by a representative of the hospital who could witness and understand our methodology firsthand. We were in dialogue with both the public and the private sector but as our main work was about users, we felt engaged in creating a good relationship with hospitals and their staff.

Performing ethnography has required clear explanations and formal approvals, especially when accessing the patients and caregivers at their homes when under the care of the hospital. Negotiating and communicating the approach has been an important part of our contribution to the project. It has also allowed us to improve our understanding of how to interact with healthcare institutions, healthcare professionals, and cross-sector partners. The designers, anthropologists, and ethnographers learned this while conducting the research itself, aided by their unique position: liaising between healthcare professionals and project partners; entering patients' homes and aligning mutual expectations and needs. Of course we had to adapt our design-anthropological language, trying to translate it in a more scientific way avoiding any misunderstandings about our objective and methods. During a meeting with doctors, we faced some challenges with their understanding of the reasons for a certain approach, more qualitatively, instead of the usage of tools like surveys. Fortunately, in the same room was the radiologist, a very empathetic professional, who had encountered qualitative methods in his career, and he helped us to translate, by making comparison with the clinical study language.

Adopting a human-centric approach means advocating for a scientific methodological *apparatus* able to understand and properly represent the life conditions of the final users, negotiating expectations with partners and adapting research methods to be as non-intrusive as possible. The final goal is to provide evidence and support evidence-driven innovations that can make a difference for healthcare ecosystems. The research focus, therefore, was to consider how homecare can support patients while contributing to a reshaping of place and the care experience. This was done by addressing the ways in which remote care systems can act:

to change the experience of the home;

to re-order the place of care-work, as new actors become enrolled within the care network and existing caregivers take on different roles and responsibilities.

**LOOKING TO THE PRESENT TO IMAGINE THE FUTURE OF ALTERNATIVE HEALTHCARE SERVICES**

Many of the companies that develop medical devices are engineering based, yet they do not integrate a user centered or participatory design approach in their development process. Most of the devices that were tested as part of the pivotal project study have a unique genealogy that considers a different kind of user than those at the center of the two homecare services we researched. For example, they might ignore the role of caregivers, or fail to consider chronic patients who are not experiencing acute episodes, or dismiss the delicate relationship between doctors, nurses, and their patients. Perhaps the technology employed was originally created for nephrological patients able to self-manage their treatments. It may not, therefore, think about elderly people with dementia. Or, it may assume caregivers are young people and not daughters and sons in their 60s who are not used to technology, or anyone who is a non-native Italian speaker. Ethnography was conducted to collect and understand user requirements, which were then shared during participatory workshops attended by the partners of the project, including representatives of the two hospital homecare services involved in the study. More than ethnography itself, these workshops gave us the chance to present the analysis derived from it, under the form of user requirements.

We began by organizing a participatory stakeholder workshop, which was an excellent opportunity to bring together key players to discuss relevant topics and engage with new ways and tools to research homecare effectively. The success of the workshop was determined by two key aspects:

1.  For the first time, partners could have a clear picture of the homecare service provided by the geriatric department thus understanding the different roles and needs of healthcare professionals, caregivers, and patients.

2.  The 25 partners involved could align themselves with each other's functions and devices (for the first time) and define where to locate their solution within the services process.

As the solutions provided by the different partners were already developed, our role was mostly to escalate some of the functions, bearing in mind that most of the devices were originally developed for other kinds of patients or clinical contexts. Most of the companies partnering on the project saw user and ethnographic research as new concepts that could be integrated in their work. Our research found that caregivers desired to have more self-reliance and assurance that they were performing care tasks correctly, and so our recommendations focused on identifying ways that these systems could provide safety mechanisms and instill confidence. Some examples of user requirements we shared with partners were: in case of wounds, it could be beneficial for both caregivers and nurses to share pictures from a distance in order to monitor the correct procedure of medication and keeping as an archive, in order to avoid loss of record keeping information between nurses. Another example was for the caregiver to consult video tutorials or other toolkits in order to manage unexpected situations as many caregivers avoid calling nurses for fear of disturbing

them. Additionally, having the possibility to communicate with CRESLA from home, through video-calling, especially when the patient is in the first phase of adaptation to the ventilator and the caregiver might not be completely comfortable with the self-management of the device. Sometimes technology has to respond to apparently small needs where the level of innovation is not on the technical aspect but on the impact on the quality of life and well-being of users.

As a result, we had two main successes during the research—three if we consider the challenging path to gain the hospitals' trust. The first success was related to the partnership between two companies: Celi, an AI and language technology company, and Panacea, a Società Cooperativa Sociale (*social cooperative organization*). The two companies decided, based on our research, to collaborate and create a virtual assistant which will be integrated into educational material for caregivers' support. This can be considered a success because their solution was specifically designed for the context of the project, personalized on the relationship between nurses and their patients, and the training sessions with caregivers.

The second success was the design of a prototype concept as a solution aimed at meeting the user requirements identified during UX research activities. A hi-fi prototype platform named GoCare was finalized, taking into consideration major pain-points and user needs—especially those regarding the management of the hospitalization at a homecare service, and the logistics planning required for its delivery. GoCare is software that enables medical staff to monitor the status of patients, schedule the home visit calendar, and manage logistics, and medical teams. It is important to understand that in homecare services, there is an externalization of hospital services to home, meaning that a complex number of people and resources are moving around the city. This means that thoughtful organization and communication is a must. Currently everything is done manually, so there is a lot of paperwork that moves in and out from the hospital. The main objective of the platform is to support organizational tasks, allowing more efficiency but also more time with patients.

Informed by ethnographic research, the digital platform is meant for doctors and nurses and is integrated with the needs expressed by the caregivers. It was presented to the medical staff and usability testing is currently in progress. The platform is receiving positive feedback considering the high value of optimizing and simplifying daily work. While it can relieve the staff from paperwork and workarounds in their flow, it is important to note that the platform is based on a highly significant element that was identified and understood only due to the research: the intimate and trusting relationship between the medical staff and their patients and caregivers. Indeed, this aspect is at the center of the homecare service and still represents the most valuable characteristic that enables hospitals to provide good (e-)care. In a world of increasingly heterogeneous and interconnected contexts, and domains of design, production and use, the aim is not merely 'getting closer' to final users (healthcare practitioners, doctors, patients, caregivers) and real-life contexts, through familiarization, mediation, and facilitation. The aim is also to create a critical and theoretically-informed distance from which to perceive and reflect upon complex and situated relations between people, technology, and design.

## CONCLUSION

The Casa nel Parco project is ongoing and will be closed by the end of the 2020 calendar year. The devices developed during the project are part of a trial, including testing their

efficiency in adapting to homecare service and reducing caregivers' stress by implementing the service. Before concluding, we want to share the most important lesson learned, which is related to how difficult it could be to enter the field and perform ethnography in hospitals. Hospital institutions are closed and very protective environments that do not speak our same language. Involving them in the research and partnering with them during the work has been the key action to create trust and collaboration. Human-centricity and participatory design are often buzzwords, especially for medical devices that may show a one-size-fits-all mentality. Ethnographic data is actionable, but deserves to be explained via workshops and participation. Ethnography has the enormous value of bringing people together. When ethnography is told, the audience is allowed to see beyond their algorithms and measurements. It helps in designing functionalities that can leverage on the strengths of human relationships. Homecare services provided by hospitals are based on the idea that they can be effective when the entire care ecosystem is human centered. These kinds of services change the relationships between doctors and patients, allow caregivers to perform an active role and delete the hostility of the clinical environment. These are all elements that can only be seen and told through the ethnography and every future technology should consider taking this lesson into account in order to start from there. In the next phases of the project we will look at caregivers' and healthcare practitioners' experiences with the devices, for new narratives to emerge. Ethnography actively contributed to conceptual debates around institution and ex-titution—that is, the de-territorialization of the physical structure of the institution and its re-manifestation through new spaces and times that seek to end interior and exterior distinctions.

**Isabel Farina** is medical anthropologist and Research Lead in the Health Business Unit in Experientia. isabel.farina@experientia.com

**Elena Messina** is medical anthropologist in the Health Business Unit at Experientia, with experience as an anthropologist working in transplant department at the Molinette Hospital of Turin. elena.messina@experientia.com

**Michele Visciola** is CEO of Experientia Global based in Basel and President and founding partner of Experientia srl based in Turin. michele.visciola@experientia.com

**Elena Guidorzi** is Research Unit Coordinator and UX Researcher at Experientia. elena.guidorzi@experientia.com

**Chiara Agamennone** is Service Design and Project Manager at Experientia in the Health Business Unit. She has a bachelor degree in nursing and worked in the public health sector from 2012 to 2016. chiara.agamennone@experientia.com

**Daria Cantù** is Senior Service Designer and Project Lead of the Health Business Unit at Experientia. daria.cantu@experientia.com

## REFERENCES CITED

Casa nel Parco (Translated as House in the Park Project), http://casanelparco-project.it/#progetto, ERDF Funds 2014-2020.

Delle Fave, Antonella & Marsicano, Sergio, 2004. "*L'umanizzazione dell'ospedale: riflessioni ed esperienze*". Milano: Franco Angeli.

Goffman, Erwin, 1961. "Asylums. Essays on the social situation of mental patients and other inmates". Anchor Books, Doubleday & Company: New York.

King, Stanley, 1962. "*Perception of illness and medical practice*". Russell Sage Foundation: New York
Milligan, Christine, Roberts, C., & Mort, M. 2011. "Telecare and older people: Who cares where"? *Social Science & Medicine* 72:347-354.

Mutch, Alistair, 2013. "Sociomateriality – acting the wrong turning?" *Information and organization* (23:1) 28-40.

Orlikowski, Wanda & Robey, Daniel. 2001. Information technology and the structuring of organizations. *Information System Research* (2:2) 143-169.

Van Der Geest, Siaak & Finkler, Kaja, 2004. Hospital ethnography: Introduction, *Social Science & Medicine*, vol. 59:10, 2004, 1995-2001.

Zaman, Shahaduz ,2004. Poverty and Violence, Frustration and Inventiveness: Hospital Ward Life in Bangladesh. *Social Science and Medicine*, 59, 2015-2036.

**THEMATIC SESSION**

# Humanizing Scale

In this session, we challenge assumptions of scale and explore what happens when we reject the notion that bigger is better and instead redefine scale to serve human values rather than corporate demands. How can scale be reconceptualized to address and meet community and human needs, first and foremost? We explore not only what scale is, but what should be scaled.

Session Curators: Bec Purser, Heli Rantavuo, Frank Romagosa

PECHAKUCHA

# Scaling Dignity

## An Antidote to Poverty?

LORENN RUSTER, *3A Institute, Australian National University*

*A wise woman once shared with me that the opposite of poverty isn't wealth. It's dignity. In a world where scale is about optimising for something bigger, faster, easier, broader and more profitable, we risk decision-making that is at odds with preserving, enabling and enhancing human dignity. What if we changed our focus to instead work out how we scale human dignity?*

*This PechaKucha draws on my career across consulting, social enterprise and academia in geographies from Sydney CBD to rural Uganda and highlights three moments where I experienced dignity that I believe can scale. Through the telling of stories it shows glimmers of how we can choose a definition of scale that preferences dignity. It can look like making space for a chicken gift, enshrining dignity in our organisational values and structures and building question-asking muscles.*

*If we believe that the opposite of poverty is dignity, then scaling dignity is an antidote to poverty. And it is within our hands to make choices that consciously value dignity when designing research processes, organisations, products and services of our collective future.*

*Keywords: Dignity, Poverty, Storytelling, Scale*

Mr & Mrs Kabi. © SolarNow

**Lorenn Ruster** (lorennruster@gmail.com) is a social justice-driven strategy consultant, intrapreneur & applied cybernetics student at Australian National University's 3A Institute, interested in the intersection of technology, cross-sector collaboration, impact, systems change & human compassion.

# Growing Communities

## How Social Platforms Can Help Community Groups Achieve the Right Scale at the Right Time

CALEN COLE, *Stripe Partners*
CAROLYN WEI, *Facebook*

*Supporting communities on its platforms has been a part of Facebook's core mission since 2017. Early understandings of the needs of groups and organizers largely centered around groups that began on Facebook itself. This paper is the result of ethnographic research conducted in 2019 to better understand the needs of different types of groups and the corresponding ways that technology platforms do and could support them. The initial orientation towards online groups led to the recognition of the difficulty of managing fast-growing groups but failed to consider whether groups might want to avoid growth in members altogether. We found in our research that many groups in fact did want to avoid or limit their growth in numbers. For these groups, growing as a community meant different things: offering more to existing members, raising awareness, or promoting the group to an outside audience, or simply maintaining over time. Our research was able to connect the dots of why organizers would have different aims between different groups or at different points in time. We ultimately presented our findings in a simple framework of three 'toolkits' that technology platforms can provide to meet the different needs of groups and organizers.*

## BACKGROUND TO THE PROJECT

In 2017, Mark Zuckerberg, Facebook CEO, announced the company would have a focus on building the "social infrastructure for community," or giving people the tools they need to build communities, a major shift from the original focus on connecting family and friends (Zuckerberg 2017). At the keynote for Facebook's F8 Developer Conference in 2019, Zuckerberg noted that Facebook is "making communities as central as friends" (Bloomberg 2019). Using the lens of community, the company has released new products and design improvements such as putting Groups, Facebook's product for people to engage with others who share their interests, front and center on the Facebook app experience. As part of this stream of work a new tab made it easier to access Groups and features were introduced to meet the needs of specific types of groups (Facebook 2019).

The work of weaving community into the Facebook app is iterative and ongoing. As such, the Facebook research and product team continues studying how people experience and build community and designing and building products that better support these behaviors. Part of the implicit challenge of this kind of large, long-term initiative is that a large product team needs to be brought along in the journey of learning about how people create and interact with their communities. Therefore, the research needs to shed light on new problem areas as well as helping the product team stakeholders internalize foundational insights that underpin multiple features. The study reported here was meant to further

progress on this overall initiative by looking at specific questions, with the findings incorporated into the general narrative understanding about community.

## ACADEMIC PERSPECTIVES ON COMMUNITY GROUPS

Academic research on community groups and organizers has primarily focused on community organizing as a response to a problem (Escandón 2010, Mundell et al. 2015) or as a theoretical process ('Community Development') that needs to be figured out and perfected by those seeking a particular outcome (Schwartz 1981). In this body of research, 'communities' are often taken for granted as homogenous pre-existing units, defined at their limit of geography or demographics.

Studies of online groups have often focused *exclusively* on virtual communities (Rheingold 2000, Wilson and Peterson 2002), while some theorists of community have sought to deny the possibility of virtual communities at all (Calhoun 1991). In our research we were particularly interested in how groups span the offline and online world and how this balance shifted depending on the group's primary orientation.

We were interested in groups that were not just responding to a problem or pre-defined by a certain geography or social unit. We wanted to explore as well the communities that are formed, joined and left in a purely voluntary manner. The community groups we encountered more closely resembled the grassroots campaigns characterized by Stokes Jones in an early EPIC paper (2005, 46) as "emergent in nature; as rooted in experiential being together; and as human projects driven by affect and effervescence as much as efficiency and purposiveness." We saw groups that were emergent projects: overlapping and crisscrossing, coming into or dropping out of existence, of decidedly uncertain long-term viability.

Goodsell and Williamson (2008) provide a wonderful case study of a hybrid community group, a group that would have fit perfectly into our study. For online communities rooted in "geographic-place-as-practice" (253) they provide the following persuasive list of what members need to sustain community: information and explanations, "hot" topics, humor, maintenance of control over interactions, mutual encouragement and connection of online and offline worlds to facilitate interaction in both (260-1). These needs resonate with our findings and their in-depth explanation of a single group focused on urban rejuvenation is enlightening. They do not, however, explore how a group's needs might change over time or how one group's needs might differ from another, particularly the differences between groups that are more offline or online oriented.

## RESEARCH OBJECTIVE

The study's goal described in this paper was to understand how people use multiple tools and techniques to create and support community. Our prior research suggested that communities cut across technology platforms. Although some communities may center on a particular digital platform like This Cat is CHONKY, a private Facebook Group for fanciers of plump cats (Kooser 2019), many other communities exist in a number of ways such as a college alumni association that can be supported by email lists, association magazines with updates about classmates, and in-person gatherings.

The study's other goal was to understand more how community organizers managed their communities across platforms and how community members engaged across

platforms. We wanted to understand how they thought of their communities and did their work with a portfolio of tools by both organizers and members. We were curious about how tools were chosen and also how digital tools supplement in-person meetings. The other area of investigation was whether the tools are a permanent part of the community's communication repertoire or if they might be gradually added, much like the communication repertoire for a person's social communication (Licoppe 2004). The product team would use this information to understand people's underlying needs for building and engaging with community and brainstorm ways to improve its portfolio of products to serve unmet needs. We had previously highlighted Facebook Groups as a place for community, and there are other features of the Facebook app like Facebook Events and Pages for businesses and organizations that could be employed.

## RESEARCH APPROACH

In light of the research questions around communities that spanned different platforms and wanting to understand both community members and organizers, we designed the study to reach a broad range of group types. Our goal was to get a variety of situations for diverse insights.

While community can be interpreted in many ways, we had a working definition: a collection of people, from which members receive a sense of belonging, connection, and feeling of safety, and to which they give trust and investment over time.

A feeling of safety means an environment that feels secure, where members understand the norms/culture/rules and how they should behave. Members can therefore do not need to worry about inadvertently breaking the rules and they can reveal more of themselves without fearing negative feedback.

Following this definition, whether a group is really a community is subjective and depends on the relationship between the individual and the group. One group member may feel strongly that it is a community for them and another may feel it is not.

### Population and Diversity of Group Types

We knew from the outset that we wanted to include a diverse range of groups. We hypothesized that three key variables would carry a considerable impact on our subject matter: the purpose of the group, the role of technology for the group and, temporality.

- *The purpose of the group:* hobby and interest-based groups, cause-based groups, values-based groups, experience-based groups
- *The role of technology:* offline-oriented vs. online-oriented vs. hybrid groups
- *Temporality:* maturity of group in its lifecycle (early-stage vs. well-established); periodicity (seasonal impact and frequency of interactions and activities)

We chose a diverse set of groups so as to uncover different and perhaps unexpected needs and opportunities.

- Religious community groups

- Neighborhood groups (city neighborhood, village groups)
- Local political groups or issue-based groups
- Hobby groups (e.g. cyclists, foodies, photographers, pet lovers)
- Women's groups, men's groups, parenting groups
- Immigrant groups
- Charity/fundraising groups

We believed these types of groups would be likely to result in rich discussions of their needs and experiences. We anticipated that they would partake in a range of kinds of interactions (e.g. in-person and online) with different membership structures (e.g. flat vs. vertical organizational structure, single group vs subgroups) and models of participation (e.g. core and peripheral members).

## Method

We chose a multi-method approach for this project as we needed to get the most out of a short five days we had together in Madrid. We ran Mobile Diaries before the in-person fieldwork to get a sense of potential participants' relationships to their groups and ensure we got the right mix of group types. We conducted in-depth interviews to capture detailed accounts of particular groups and the needs and descriptions of the tools used by each group. The team partook in observational immersions of community spaces and group meetings to get a feel for the cultural context and the emotional energy that these spaces and activities engender. Lastly we conducted brief unscheduled intercept interviews with people we approached in public spaces in order to get signal on an even wider array of groups and let us compare our specific sample with random representatives of the 'general population' so we could see if our respondents seemed to be outliers in their level of group engagement. Intercepts and immersions also gave us optionality, allowing us to take advantage of spare blocks of time as we felt best suited us in the moment.

## KEY FINDINGS

## The Role of Offline

Because of the history and culture of Facebook, the nature of online groups and the examples of online communities were always prominent in conversations. It was essential that we learn more about offline-oriented groups to understand how they differed from online-oriented groups and the different role that technology might play for them.

In our interviews, we asked participants to map the groups that were most important to them. Time and time again we saw the same groups closest to the stick figures representing our respondents: *la familia, amigos/as de toda la vida, amigos/as del pueblo:* family, lifelong friends, friends from my hometown. Many residents of Madrid are originally from small towns, and we often heard about the importance of *pueblos*, with town associations and annual parties, as a place to return to and be around one's family and lifelong friends. When we asked about tools and platforms, these groups always 'lived' digitally in one place or another, most often

as groups on WhatsApp. But while tools kept members of the group in touch, the groups existed in their 'pure' form outside and beyond any technology platform.

Immersions to community spaces showed us key differences between online and offline spaces for groups. We visited a cultural center run by local community groups and volunteers. The former tobacco factory had been transformed into spaces for lessons, rehearsals, and performances of dance, and music, a community garden, craft workshops and studios for sculpture and painting. A large calendar listed out the upcoming events and their locations while a bulletin board overflowed with flyers for events, classes, services and resources.

In online spaces, groups typically stand on their own as discrete and disconnected units. The cultural center in contrast was defined by a space of overlapping and intermingling groups. Regulars could drift from one activity, group or space to another. Non-members could come without a clear goal in mind to peruse the activities, observe groups in action and mingle with the other humans serendipitously sharing the space.

As we learned about groups which were oriented around offline activities, we were surprised to see that their interactions on digital platforms were often devoid of purely social light-hearted or humorous communication. There was little interaction that demonstrated intimacy and trust between members of the group. We realized that they had much less need for these types of interactions when online because their in-person activities provided the sufficient and appropriate time to bond and enjoy each other's presence. In contrast, the conversation on platforms was nearly all sorting logistics and sharing necessary information to achieve their in-person goals.

## The Role of Growth

As discussed above, teams at Facebook were very familiar with online groups, which could quickly grow from zero to hundreds of thousands of members. We'd seen how this could bring challenges for admins in moderating and managing such groups. We knew that not all organizers were trying to grow their groups all the time, but we didn't have a clear understanding of why or why not. Nor had we thought of growth as a potential negative, to be actively avoided.

As we gained a deeper understanding of the diversity of groups that our participants were a part of, many revolving around offline interaction, we learned that many were quite content with the number of members they had and sought no growth or only limited growth in numbers. For some groups, an in-person activity at the center of the group limited from a logistical perspective the number of participants. This was the case with groups that came together to play sports. For other groups, the pool of potential participants was limited by characteristics required for the relevance of the group to a given individual, for example a group for immigrants from one country living in a particular part of Madrid. In a third scenario, some groups put tight limits on growth to ensure the quality and safety of their members. This was the case with a volunteering organization that had an extensive vetting process, requiring interested candidates to apply, attend an in-person meeting and run a trial of the service they intended on offering.

It was not only offline-oriented groups that were content without growth. We encountered online-based groups for sharing common interests, whether in collectible figurines or BMW 7-series, where organizers and members felt that no need for new

members given that their groups regularly had members posting content and communication put out to the group resulted in engagement and feedback. Adding new members at this point would simply increase the likelihood of irrelevant or inappropriate content.

There was a handful of groups that were actively engaged in increasing the number of members. These were all groups that were centered around a particular cause or agenda. One group wanted to promote a particular philosophy within medical practice. Another was comprised of parents organizing to promote education for children with autism while another sought to improve the working conditions of prison guards. These groups were all hybrid groups, with significant activities both offline and online. The former was often used for planning and decision making, while the latter was essential for reaching new audiences.

In general, provided there were enough members to make the core activity possible and to keep the group active and lively, online or offline, these groups did not seek growth for growth's sake. It was only with groups where the core purpose of the group depended on increasing the number of members or reach of the group that gaining new members and raising awareness was a top priority.

## Growing beyond Numbers

Sheryl A. Kunjawa-Holbrook (2017, 203) writes of groups partaking in interreligious learning that:

> Some congregations experienced growth beyond numbers-profound spiritual growth, growth in the knowledge of their own tradition, growth in community involvement, growth in hospitality, and growth in relationships between members due to interreligious partnerships.

We saw many organizers striving after a similar 'growth beyond numbers'. Groups like a soccer team and a collection of dancers were limited in the number of participants, but as they participated in exhibitive activities they did seek to promote a community of fans and drive attendance at performances.

Others sought to grow the richness of the experience, finding new sources of common ground with other members. One woman combined her love of a photography app with her hobby of collecting dolls by creating a subgroup of fellow members who shared both passions.

We met organizers who wanted to grow their groups in stature more than in numbers. Andrea had started a Facebook group to donate items to people in need in his neighborhood. At the time of research, the group had nearly 1,500 members, and Andrea had 13 volunteers working with him yet his primary goal was to obtain official status as a non-profit, a permanent physical location, and some full-time staff members.

## IMPACT

## A Framework for Technology Supporting Community Groups

After our research, we faced a large number of stakeholders across products and roles keen to learn about what community groups needed and how technology platforms could support them.

The findings had to be actionable – the ultimate objective was always creating a platform that better supported community groups. They had to be communicated in a way that was clear and consistent: with a large audience we couldn't risk conflicting interpretations that could arise from an overly complicated analysis.

We also needed to present our findings in such a way that they would instantly gel with intuitive common sense and personal experience, for two reasons. Firstly, nearly everyone has some experience being a part of a community group. With such a topic a listener will naturally compare the findings of the research with her personal experience. If the two clash, the findings are likely to be regarded with uncertainty.

Second, was the scale of the internal initiative and the broad group of stakeholders. Unlike a product team working on a sprint, the initiative to support communities is a long-term strategy and mission across many different teams and products at Facebook. For our findings to have an impact, we needed the learnings to embed with many stakeholders such that it would stick with them over time. Our findings needed to be easy to understand without too much effort and they needed to be not just believable but *deeply believed* by our audience.

Although we needed actionable insights, the level of specific features and UI was going to be too granular. We had to think about what it was that organizers and members were trying to achieve. Communication was at the heart of what platforms provided to groups, but to what end?

Ultimately we formulated a framework of three broadly scoped 'toolkits' through which a technological platform could meet the different needs of community groups:

- Managing logistics
- Promoting the group or a cause
- Facilitating discussion

*Managing logistics*: Who is doing what? What are the key dates and deadlines? Who is bringing what? Who has paid and who hasn't? Who has an extra seat in their car?

*Promoting the group or a cause:* Adding new members, fundraising or collecting donations, raising awareness

*Facilitating discussion:* Discussion that is active, a responsive community where members answer each other's questions, give advice and make suggestions, sharing ideas and inspiration, enjoying relationships and making new connections when not together in-person. Not too quiet and not too noisy.

All groups and organizers at certain points will have a need for all three toolkits. However, a group often has a much stronger need for one toolkit over the others, and further, the relative importance of each can change considerably depending on the stage of the group's development.

## A Platform that Evolves with the Group

Concrete understanding of these groups helped the product team refine its definition of community, with a greater focus on the purposes and blurred boundaries of communities. Influenced by this study, the team shifted from thinking about individual features to a portfolio approach to the features for community organizers, recognizing the need for toolkits of features that serve community organizers' various goals and needs. The growth of the group is only one need served among many others. Further, the team began considering the development stage of a community as a factor that influences priorities. Most groups tend to start small. We saw how groups evolved and how leaders evolved with their groups. We understood that toolkits provided must be similarly flexible and able to evolve with a given group.

The existing body of knowledge within Facebook was enriched by having a much strong appreciation of the notion of life stages in a community and in each person's relationship or journey with the community – and thus differing needs for toolkits over time. The insights from the study helped frame the team's community strategy, which includes yet to be released products that better support the different life stages of where a person might be in their journey to community, for example helping users discover communities of potential interest at the beginning of a journey or further along helping them engage more deeply in communities they already belong to.

From a researcher's perspective, the study confirmed a hunch that we needed to think more holistically about the goals of community leaders. Although efficiency and growth are important at different times, considering their deeper motivations and aspirations for the community can help us provide a richer set of tools. With the themes of the toolkits, we have been shifting the narrative about leaders to one of longer-term goals and not just short-term efficiency. Likewise, with the more nuanced understanding around the desire for growth or lack of, we can consider other tools for preserving internal culture and nurturing.

From a storytelling perspective, we were able to tell more stories about the goals and vision of community organizers. More tactical research would highlight a narrow problem that Groups admins might experience (such as working through pending membership requests) and suggest possible fixes. This ethnographic fieldwork enabled us to discuss the bigger picture of goals and vision – what organizers were truly concerned about (often preserving culture rather than growing the group). We were then able to place the tactical problems within the context of broader goals to explain why they were problematic in the first place (for example sifting through membership requests from people who are not the target audience of the group doesn't help achieve the ultimate goal of preserving group culture).

## Equal Appreciation of Online and Offline

Previous research was heavily grounded on understanding around communities based on the Facebook platform. Our insights added much-needed clarity into group dynamics, functioning and priorities in more offline-oriented or hybrid offline-online contexts, areas that as discussed were less well-known to the team. We were able to explain to stakeholders why an organizer might use a mix of tools, whether a multi-purpose communication tool like

WhatsApp or something perceived to be ideal for certain situations such as sharing photos on Instagram.

The sorting of offline logistics in particular showed that a good deal of the interactions that take place on digital platforms are about meeting basic needs and ensuring awareness between members, whether coordinating a potluck, organizing a gift exchange or finding a carpool to this week's soccer practice. Fewer interactions are actually *about* the shared interest or purpose of the group. This insight set the stage for a design sprint exploring how we could better serve the unique needs of community leaders and inspired further exploration of how teams of community leaders coordinate and communicate amongst themselves. Ultimately the sprint gave specific product teams some action items to develop features that can help leaders be more proactive in meeting their vision for their community.

The vision of supporting communities underpins much of the product and touches many different teams and features. We generated knowledge with this research which was adopted and internalized by a huge team, ultimately making the research a success.

## DISCUSSION: LEARNINGS FOR THE EPIC COMMUNITY

### 'Obvious' Findings and the Scaling of Findings to Large Numbers of Internal Stakeholders

The organizational psychologist Adam Grant has written about the value of 'obvious' insights (2019). Findings that resonate with common sense and intuitively 'sound right' to the audience can, Grant argues, be particularly effective in overcoming three barriers to change: resistance to new data, resistance to change and organizational uniqueness bias.

> Obvious insights can motivate us to close the knowing-doing gap. Common sense is rarely common practice. If you ask managers what effectiveness looks like, they often can spell out the critical factors. The key is to get them to act on that insight, and that's where the obvious can help.

That community group's goals and needs change over time and depending on the group's purpose is not particularly surprising or particularly profound. In our case, we believe it was an advantage rather than a weakness to present findings in a simple framework that was compatible with common sense experiences and well-established understanding of what community groups were like, rather than challenging conventional wisdom or introducing unfamiliar concepts.

As researchers, we often feel the desire or need for our findings to unveil a new paradigm, which overturns the previous understanding of the subject matter, shatters preconceived notions, and busts common-sense myths. In this work, we had to resist such temptations or risk the findings being taken opposing directions or falling by the wayside entirely.

In presenting a straightforward framework that was easy to grasp intuitively, we sought to provide a solid foundation of shared understanding that would serve as the basis for decisions and execution over the long-term across a multitude of different teams and products. It is our belief that this approach is likely the more effective one with research projects seeking to drive long-term impact with many diverse stakeholders, particularly when

the initiative is an ongoing one touching on multiple parts of the product (as opposed to, for example, a brand new product being launched).

## CONCLUSION

The goal of this research was to understand the needs of community groups and group organizers and the role of tech platforms to inform the organization's strategy of supporting community groups, ultimately seeking to embed insights with diverse roles on a multitude of different product teams. We learned about a range of group types, in particular gaining knowledge about offline-oriented and hybrid online-offline groups.

In our research we saw that growing the number of members was only a priority in a select number of groups. Others sought to limit or even actively avoid too much growth. They sought instead to grow by increasing the richness of the experience of their members by increasing the stature of the group or by increasing engagement with an outside audience.

We identified the three main 'toolkits' that community groups need from technology. Most instructive for our stakeholders was understanding how these needs varied depending on the purpose of the group and on the stage in development of the group.

Communicating our findings in a simple and straightforward framework that fit intuitively with stakeholders' common sense and avoided abstract or unfamiliar concepts was key to the impact of the research. This framework along, with concrete examples from case studies, allowed us to embed a large number of stakeholders on a large number of teams with a shared foundational understanding of what community groups need from technology that would enable them to make decisions about product and feature development as this initiative continues over the years.

## REFERENCES CITED

Bloomberg. 2019. "Facebook leans harder into Groups as it redesigns main app." LA Times website, April 30. Accessed September 22, 2020. https://www.latimes.com/business/la-fi-tn-facebook-redesign-f8-groups-oculus-zuckerberg-20190430-story.html

Calhoun, Craig. 1991. "Indirect Relationships and Imagined Communities: Large-Scale Social Integration and the Transformation of Everyday Life." In Social Theory for a Changing Society, edited by P. Bourdieu and J. S. Colemen, 95-121. Boulder: Westview Press.

Escandón, Socorro. 2010. "Theoretical Versus Grass-Roots Development of a Community Partnership." The Qualitative Report 15 (1): 142–155.

Facebook. 2019. "Day 1 of F8 2019: Building New Products and Features for a Privacy-Focused Social Platform." FB.com, April 30. Accessed September 22, 2020. https://about.fb.com/news/2019/04/f8-2019-day-1/

Goodsell, Todd L. and Owen Williamson. 2008. "The Case of the Brick Huggers: The Practice of an Online Community." City & Community 7 (3): 251-71.

Grant, Adam. 2019. "The Surprising Value of Obvious Insights." MIT Sloan Management Review, February 5. Accessed September 22, 2020. https://sloanreview.mit.edu/article/the-surprising-value-of-obvious-insights/

Jones, Stokes. 2005. "Grass Roots Campaigning as Elective Sociality (or Maffesoli Meets 'Social Software'): Lessons from the BBC iCan Project." Ethnographic Praxis in Industry Conference Proceedings: pp. 31–52.

Kooser, Amanda. 2019. "This Cat is Chonky: The fat-cat online shrine lifting humans from despair." Cnet.com, June 14. Accessed September 22, 2020. https://www.cnet.com/news/this-cat-is-chonky-the-fat-cat-online-shrine-lifting-humans-from-despair/

Kujawa-Holbrook, Sheryl A. 2017. "'The Sacrament of Human Life' Cultivating Intentional Interreligious Learning in Congregations." In Teaching for a Multifaith World, edited by E. S. Fernandez, 199-217. Eugene: Pickwick Publications.

Licoppe, Christian. 2004. "'Connected' Presence: The Emergence of a New Repertoire for Managing Social Relationships in a Changing Communication Technoscape." Environment and Planning D: Society and Space 22 (1): 135-56.

Mundell, Leah, Lisa J. Hardy, Roxana Cardiel De Niz, and Michelle Thomas. 2015. "Defying Displacement: Organizing for a Beautiful Life in Flagstaff, Arizona." Practicing Anthropology 37 (2): 35–39.

Rheingold, Howard. 2000. The virtual community: homesteading on the electronic frontier. Cambridge MA: MIT Press.

Schwartz, Norman B. 1981. "Anthropological Views of Community and Community Development." Human Organization 40 (4): 313–322.

Wilson, Samuel M. and Leighton C. Peterson. 2002. "The Anthropology of Online Communities." Annual Review of Anthropology, Vol. 31, 449-467.

Zuckerberg, Mark. 2017. "Building Global Community." Facebook.com, February 16. Accessed September 22, 2020. https://www.facebook.com/notes/mark-zuckerberg/building-global-community/10154544292806634/

CATALYST

# Where Can We Find an Ethics for Scale?

## How to Define an Ethical Infrastructure for the Development of Future Technologies at Global Scale

IAN DULL, *ReD Associates*
FANI NTAVELOU BAUM, *ReD Associates*
THOMAS HUGHES, *ReD Associates*

*Despite companies facing real consequences for getting ethics wrong, basic ethical questions in emerging technologies remain unresolved. Companies have begun trying to answer these tough questions, but their techniques are often hindered by the classical approach of moral philosophy and ethics – namely normative philosophy – which prescribe an approach to resolving ethical dilemmas from the outset, based on assumed moral truths. In contrast, we propose that a key foundation for 'getting ethics right' is to do the opposite: to discover them, by going out into the world to study how relevant people resolve similar ethical dilemmas in their daily lives – a project we term 'grounded ethics'. Building from Durkheim's theory of moral facts and more recent developments in the anthropology of morals and ethics, this paper explores the methods and theory useful to such a mission – synthesizing these into a framework to guide future 'grounded ethics' practice.*

*Keywords: Ethics, Technology, Methodology, Moral Facts*

## INTRODUCTION: THE ETHICS FRONTIER IN TECHNOLOGY

Debates around the moral permissibility and the ethical implications of technology have long been ongoing in global fora (Moss and Metcalf 2019). In recent years, as technology has become a foundation for everyday life globally, such debate has become commonplace in public, academic, and practitioner fora as part of broader 'tech-lash,' leading to its questioning in political discourse and hearings, frequent publications in major news organizations, academic fora including EPIC (see Tamminen and Holmgren 2016; Moss and Schüür 2019), and discourses within the technology industry and organizations – what technology news outlet *The Verge* has termed, "an ethics explosion" (Vincent 2019).

This renewed attention has centered on a growing number of issues often in response to scandals around new technology practices, including but not limited to: algorithmic bias (i.e. in mortgage reviews or prison sentencing, see Eubanks 2020; Lee et al. 2019; Anderson et al. 2019), data collection (i.e. surrounding the Cambridge Analytica scandal or location tracking), privacy (i.e. the 'right to be forgotten'), freedom of speech, and exploitative practices (i.e. surrounding Uber and the 'gig economy'). Furthermore, momentum sparked by such debates has begun to spill over into discussions of future applications of new technologies, including augmented and virtual reality (AR/VR), more generalized artificial intelligence, self-driving technology, and brain-machine interfaces (i.e. Neuralink). With many major technology players becoming embroiled in unforeseen ethical controversies as a result, the industry is now increasingly concerned about the ethical implications of product

decisions and future technology investments, in terms of future product usage, adoption of new products & technologies, brand equity, and even employee satisfaction.[1] Recent commentators have noted how this has even begun to enter the language of corporate financial risk: Metcalf et al. (2019, 451) note how Alphabet, parent company of Google and Waymo, in 2018 reported to investors that AI products could "raise new or exacerbate existing ethical, technological, legal and other challenges, which may … adversely affect our revenues and operating results." That many of these companies and their technologies operate at (and require) enormous, often global, scale has also meant an amplification of concern around their potential moral effects. The bigger the impact, the higher the ethical stakes.

While concerns about the moral permissibility and ethics of technology development are notable, thus far, calls for more ethical practices have too often stopped there. Though this no doubt bears reference to the scale of change being demanded of technology players, it nevertheless leaves a gap for technology practitioners who seek to heed calls for change and implement more ethical practices. As such, this paper aims to argue not *if* ethics are needed, but *where and how* they can be studied. This paper will present a framework for a 'grounded' approach, employing anthropological theory and methodology for this pursuit.

However, the pursuit of understanding the ethics of new technological development is by no means straightforward. To take just one example: when introducing new products and investing in new technologies and applications, technology companies often take guidance from the legal limits currently in force. Yet just as new technologies often engender new behaviors, entering new technological terrains also means entering new moral and ethical ones. Facebook's Live product, for example – launched in 2016 to enable live streaming of content to Facebook 'friends' and a wider audience – quickly led to freedom of speech and censorship concerns, specifically around the broadcasting of violent and terrorist acts (Issac & Ember 2016). Facebook's standards likely met legal expectations prior to Live's launch – but what of the social limits around what *should* be shared freely and without censorship, and the legal limits that may be drawn going forward? Today's legal (and moral) acceptance offers uncertain guidance for the boundaries of future moral outrage.

With a new generation of cultural-paradigm-shifting technologies on the horizon, technology companies need better tools to understand what moral standards consumers and society will hold them to – and where boundaries are likely to fall in the future. As a foray into this domain, we propose a project of 'grounded ethics' to help companies understand users' moral intuitions: When do technologies risk overstepping people's moral boundaries, their sense of right and wrong and result in public reproach, i.e. with Facebook's Live or Google's Glass? How do we locate people's boundaries around technologies like voice recognition, in light of social processes of normalization? What ethics should be scaled into agentive products themselves, like algorithm-driven feeds? How much can moral boundaries truly be global, and, if not, where should the lines be drawn? While morality and ethics have traditionally been the domain of philosophers and deduction, we argue that it is precisely the anthropologist's inductive approach that is needed to develop the problem-specific and practical frameworks needed to guide the executives, technologists, and designers who will make future decisions. Building on Durkheim's concept of the moral fact and recent work in the anthropology of morals (see Fassin 2006, 2014), we aim to propose an initial framework for how to study the moral landscapes surrounding future technology applications to create the ethical infrastructures for future technology development.

## BUT WHICH ETHICS? A THEORETICAL CHALLENGE FOR THE PRACTICE OF SCALABLE ETHICS

Calls for more ethical approaches to the development and deployment of new technologies have largely focused on what technology companies *should* and *should not* do. This is to say that the primary calls for more ethical action have largely been – as has long been standard in moral philosophy and ethics, and thus, society at large – *normative*: they tend to prescribe how the future behavior of technology companies should look, and suggest the rules or 'tests' according to which they should function, or consequences which they should not permit. General business ethics have tended to follow the same path, with the additional step of aligning business goals (in this case, often considered the 'norms') to the individual behaviors of employees (see Sims 1991).

To their credit, major technology firms like those of Silicon Valley have followed suit – and have begun to respond in myriad ways. As Metcalf et al. (2019) have documented, 'ethics owners' have proliferated in Silicon Valley companies – ranging from individuals responsible for developing ethical procedures and supervising teams to ensure ethical practices to even making corporate ethics a personal mission. Others, like Google, have famously hired 'in-house philosophers' to handle philosophical quandaries that may come up in the development of its products and technologies (VentureBeat 2011). These personnel-driven solutions have come in addition to more traditional and longer-standing practices, like the following of local and international regulations for safety and ethical testing, and the now common practice of consumer surveys, A/B testing, and other large-scale quantitative instruments to determine consumer interest and drive more engagement or use. Work has continued outside of corporations themselves, with universities mandating ethics for software engineers and computer scientists (Fiesler 2018), and groups of practitioners and scientists calling for codes of ethics, statements of principles, or bans on certain practices, notably the development of autonomous weapons (see Sample 2018; Future of Life 2015).

Such reinvigorated practices are relatively new within major technology firms, making their efficacy uncertain in the short-term. Nevertheless, the continued appearance of new ethical challenges to technology companies' products and practices – most recently surrounding the censorship of 'fake news' and misinformation during the 2020 COVID-19 pandemic and US Presidential Election (see Frenkel et al. 2020; Warzel 2020a) – points to significant opportunities for improvement. These discussions are all the more urgent against the backdrop of Silicon Valley's increasingly outsize influence in shaping the public sphere – the gatekeepers to how society accesses and experiences information –and even more so, when contrasted with increasingly limited (or ineffective) government and civil society institutions. Even when major tech companies have taken proactive stances on ethical issues – like Facebook's announcements around political ads in the 2020 US Presidential Election (see Isaac 2020; Warzel 2020b) – many have observed a tacit signaling of their increasing control. (And, we might add, a tacit acknowledgement that their products themselves adopt an ethical stance one way or the other). As one journalist wrote of a similar Facebook pledge in Germany's 2017 elections: "It's a declaration that Facebook is assuming a level of power at once of the state and beyond it" (Read 2017). The combination of technology players' growing necessity, power, and inconsistent performance has only reinforced initial calls for more ethical actions.

Beyond continued demands for ethical accountability, commentators have also observed many challenges facing the practice of ethics within technology companies. While ethics remains a buzz word throughout Silicon Valley, Metcalf et al. note in their study of 'ethics owners' that many everyday practitioners – the designers, engineers, managers, executives, and more who make or drive product decisions – "'are not yet moved by ethics'" (2019, 453). In more plain terms, this is to say that ethics do not enter consciously into the day-to-day practices of product development. Moreover, the authors note that, in such climates, ethics owners' mandates (vis-à-vis compliance, CSR, and others) and their roles organizationally (i.e. who they report to, how they can influence projects) are unclear, often leaving ethical concerns dangling within organizations. Even when voiced, ethical qualms are furthermore drowned out by common Silicon Valley discourses like 'technological solutionism' – the belief that *better* technology will resolve ethical problems – or 'market fundamentalism' – the belief that market indicators trump ethical decisions or that consumer demand (i.e. continued use) proves moral acceptability – have a tendency to downplay or entirely undermine otherwise legitimate concerns. In this context, weighing complex ethical decisions becomes "doing ethics" (2019, 453) – yet another task to be performed in the course of product development. That, they argue, points to a challenge as to if ethics can coexist within the current structures and internal logics driving firms.

We agree; yet these challenges represent only the organizational dimensions of the practice of ethics in technology companies. Conspicuously lacking from discussions has been the question of *which ethics* technology companies should abide by to begin with. This is to say: when technology companies bring in corporate philosophers, appoint ethics owners, or create ethics boards, which systems of ethics should they bring with them or judge proposed projects and products against?

From that perspective, calls for ethical accountability in technology companies have been quite unspecific – and it is here where a *normative* approach to defining ethics can fall short. The choice between normative ethical systems – e.g. between utilitarianism, Kantian ethics, care ethics, virtue ethics – ironically leaves open how ethical quandaries are to be interpreted and resolved, and does so in the absence of input from the people ethical decisions will affect. That challenge is not only theoretical, but empirical: normative standards for defining ethics have failed to deliver meaningful guidance on moral permissibility and ethical action, notably on three fronts:

1. *Lack of consistency:* While individual corporations have attempted to define their own normative ethics to guide corporate behavior, when looking across the technology sector, these individualized approaches to normative ethics have created different and competing systems – yielding anarchy, rather than a consistent approach to ethics. Recent research into 84 AI ethics guidelines from companies and organizations around the world found that "no single ethical principle appeared to be common to the entire corpus of documents" (Jobin et al. 2019). When each company selects their own normative moral foundations and ethical principles, as opposed to deriving them from prevailing moral and ethical tides, they are contributing to an overall climate of ambiguity that ultimately undermines the project of an ethics of technology development in the first place (see D'Ignazio and Klein 2019).

2. *Lack of nuance & context-specificity:* As an approach founded on *a priori* truths, normative ethics tends to categorical assertions, and technology is no exception – whether aspiring to full transparency with consumers, asserting or denying the primacy of privacy, or defining what tasks machines should and should not take on. In practice, few morals operate in such black-and-white terms. Recent ReD technology studies have explored the boundaries of what types of data collection can be acceptable. Many informants were unaffected by their data being collected – surprisingly even for what one might consider 'sensitive' data, like home addresses. Yet when faced with unexpected voice or video data collection – like a dubious beep from an Amazon Echo during a private, political conversation at home, or unexpected filming in public – reactions were visceral, and anger immediate. In that case, 'privacy' was not so much an absolute value, but a contextual one. Without the right qualifiers in place, normative principles can be controversial or counterproductive to commercial aims. Finding the right context and execution for a technology can drastically modify its moral and ethical dynamics.

3. *Lack of future-proofing:* In asserting one way to understand the morality of the world they occupy, technology firms' normative assessments of ethics fail to capture the shifting nature of moral systems, or account for how the technology they produce can shape moral systems. This can work to both the benefit and harm of companies' ambitions. To take a positive example, the past decades have created a major shift in public intuition around 'strangers online' – from dangerous to 50 million people on Tinder. Had the creators of Tinder only followed dominant moral codes surrounding the early internet, they might not have found the same success. Yet cautionary tales also abound: While most photography was accepted and prevalent in the early smartphone era, Google's Glass overstepped these boundaries by turning glasses – and by extension, the body – into a camera, therein reimagining norms around privacy. A meaningful picture of moral and ethical future action does not necessarily emerge from the standards in front of us today.

Taken together, these challenges point to the societal and corporate risks which a plethora of *normative* assessments of ethics in technology development can create. So why do companies still run these risks, especially after investing time and resources to develop their products and new moonshot technologies? This is not due to a lack of effort, but due to a methodological fallacy. There is a centuries-long tradition of armchair, top-down ethics: philosophers – and now corporate philosophers – have sat around, thinking about the right and wrong ways to live, based on virtues, the consequences of our actions, or the deontological imperative. But they tell only half of the story. Normative, top-down ethics has given a multitude of rulebooks for how one *should live*, but it does not say much about *how we do live*. Just because we know that lying is wrong, it does not mean that we do not lie. And as cases like the Milgram experiments clearly have proven, people rarely meet even their own standards of virtuous life. Moral philosophers have been long baffled and divided on how to trace these moral facts – and the many others – in society and how seriously to take them. As the above goes to show, the complexity and stakes of ethics decisions are too high for individual stances on morality.

Yet with the right tools, the picture can become simpler: rather than eliding the realities of moral facts in society to describe what we *ought to* do (as moral philosophers have), we suggest a knowledge of *how* people live morally and what they *will* and *will not* accept as the basis for defining an ethics for technology development. Just as moral philosophers likely should not look into absolute moral truths in lived daily life, corporate ethicists should not look for moral facts in the theoretical realm.

## FROM THE ARMCHAIR TO THE BAZAAR: 'GROUNDING' ETHICS IN LIVED MORAL FACTS

We have traced the cause for these risks down to the method of defining ethics itself. Companies tend to *theorize* what people think is ethical instead of *discovering* how ethics are navigated. In this paper, we propose an alternative approach that avoids these risks – by defining a 'grounded ethics,' designed to study and understand the nuances around 'moral facts' that govern the aspects of life a technology could change.

Durkheim still looms large in any discussion of moral facts. We embrace his view of ethics as grounded in social life, facts to be discovered through how people think and behave. While temporally far from current debates around the ethics of technology, Durkheim's theory arose in the context of the social upheaval accompanying the industrial revolution (Laidlaw 2017). As such, his theory is attuned to understanding what is moral and ethical as both 1) defined by the realities of the social world (as uncoupled from normative, religious mores), and 2) flexible with regards to social changes, for example, those shaped by new technologies. As Durkheim proposed, we suggest 'grounding' a development of ethics in uncovering 'moral facts': the pillars people use to shape a sense of living a morally good life – which may be observed and studied in culture (Durkheim [1924] 2010). In broad strokes, in his view, ethics are observable through sanctions – social consequences to rule-violating actions. The actions that would trigger these sanctions delineate what is morally acceptable and what is not. The upside of adopting Durkheim's view is that it clearly points to a domain of study and observation – the social rules that are observed and the sanctions to which they give rise.[2]

Anthropology has turned away from Durkheim's moral facts in the past, due to some commentators' interpretations of Durkheim's focus on social sanctions as representing overly fixed norms (Laidlaw 2014), or norms *simpliciter*. Didier Fassin, however, offers a helpful argument to reposition Durkheim's understanding of social sanctions to also include complex individual negotiations:

> Durkheim himself had a more sophisticated and somewhat ambiguous theory than what is often simplified by commentators, including Laidlaw (2014, 21), who writes that the French sociologist "ended up with a conception of morality as thorough law-like as Kant's, but with obedience to the law naturalized into the smooth functioning of a well-engineered mechanical system," thus ignoring what Durkheim ([1924] 2010, 17) clearly asserts: "In opposition to Kant, we shall show that the notion of duty does not exhaust the concept of morality," since "to become the agents of an act it must interest our sensibility to a certain extent and appear to us, in some way, desirable." Such an act "cannot be accomplished without effort and self-constraint" and "is not achieved without difficulty and inner-conflict" — a

language not so far removed from the contemporary anthropology of ethics.
(Fassin 2014, 430)

In this light, Durkheim's notion of a moral fact asserts not only a distinction of social norms from Kant's norms *simpliciter* – thereby creating space for cultural, individual, and temporal variation – but also locates discovering variations, and their future directions, in individual experience.  The grounded approach to ethics we propose is built on this Durkheimian proposition that moral facts are to be discovered in the lived reality of human life: in the daily behaviors and choices of individuals, the symbols they respond to, and sanctions they recognize, as they navigate towards the right or wrong side of virtue. We furthermore believe that identifying these moral facts is at its most feasible and productive when it is focused on the individual experience of choice and conflict. Ideally, this approach would be supplemented with a larger understanding of the historical factors that give rise to the norms that constrain ethical action by creating implicit and explicit sanctions.[3] We believe, however, that a focus on the individual experience of moral decision-making is more valuable to building an understanding of the moral boundaries that are likely to shape future technology products. In the same way that learning about traffic laws does not teach us everything we need to know about acceptable driving conduct, understanding social norms does not tell the full story about acceptable moral behavior.

To put our stance succinctly, we are describing a grounded ethics framework with three necessary features:

1. *It is bottom-up.* We are interested in understanding the nuances, shortcuts, trade-offs and irregularities in how people experience the moral systems they inherit and create.
2. *It assumes a scale of flexibility of moral facts.* We believe that moral facts and malleable and subject to change by the same forces that forged them in the first place, be it social, political and religious factors, tradition or biases of moral psychology.
3. *It is application-dependent.* Finally, some housekeeping. This framework is not developed with the intention to be applied to the moral character of a society or group at large. This is too big of a project and not helpful for the purposes we have in mind. Rather, we are operating under the assumption that we can secure depth and nuance by focusing on the social phenomena a given technology has the potential to transform.

While this latter proposition could be questioned on the grounds of being too narrow – how could we understand moral obligations around privacy, without understanding the context of morality more generally? – given the resource constraints placed on practitioners outside the academy, we view undertaking such complete studies of morals and ethics to be too ambitious to be practical. Rather, as we will detail in the process of outlining a research approach for developing a 'grounded ethics,' existing ethnographies of relevant societies need to suffice to provide the moral backgrounds against which more focused questions of technology applications can be studied.

## DISCOVERING A 'GROUNDED ETHICS' IN PRACTICE – A FRAMEWORK

A 'grounded ethics' is then 1) the most productive approach to identifying the moral foundations to guide the development and deployment of technologies at scale, and 2) clearly grounded in a society's moral facts, but especially in the daily behaviors, choices, and trade-offs faced by individuals living in those societies. How then should we as practitioners working with technology companies practically seek out and *discover* a 'grounded ethics,' for the real-world technology problems we are likely to face?

At this phase in applied social science research and 'UX,' a range of tools – from product-centered ethnography to usability observations and attitudinal/behavioral surveys – would normally be seen as the defaults for exploring new product and technology innovation challenges. Yet, as Amirebrahimi (2016) has already discussed at length in the EPIC forum, while these methods have proven successful at identifying new commercial opportunities through observed and emerging behaviors, promising attitudes, and a willingness to adopt or pay, such approaches have come to be co-opted and oversimplified in practice – too much so to address the "difficulty and inner-conflict" (Fassin 2014, 430) that accompany moral negotiations. To quote one of Amirebrahimi's 'UXer' informants, these methods too often "don't get at the very real issues" (Amirebrahimi 2016, 87) and by Amirebrahimi's account reduces lives into "only [a person's] moment of use" (2016, 89). To combine this critique with Metcalf et al.'s critique of 'doing ethics' (2019) would suggest that using UX approaches reduces the complex moral choices of individuals and their societies to a simple review of their "moment of acceptance" of a new product or technology – devoid of the context(s) in which such acceptance may occur, the moral 'costs' or 'burdens' of such decisions, and how flexible such moral facts are for people. That leaves the nuances of moral facts quickly reduced to binary – yes / no – permissibility.

In line with this approach, we believe that any study of 'grounded ethics' for technology development at scale must deeply explore several layers: (a) the cultural foundations of the targeted societies, (b) the 'virtuous' phenomena likely affected by the new technology, (c) the ethical interests of different social groups, and who is the moral 'user' in each case and (d) contexts relevant to *where* ethics may be applied, e.g. physical sites, varied social groups and (e) moral notions around monetization. Across these, we suggest that the foundations for defining a 'grounded ethics' for new technologies lie in understanding the social systems, moral intuitions and dilemmas, and visceral reactions around the underlying social phenomena a given technology has the power to shape. We suggest that uncovering these foundations will require incorporating methods beyond the conventional applied social science toolbox, like social experiments and experimental philosophy. Our hope is that this framework will be useful in rendering a prescriptive picture of the moral landscapes in which companies balance ethical trade-offs.

### Cultural Foundations

As many anthropological studies of morality and ethics (see Laidlaw 2014; Widlok 2004) have made clear, the morals and virtues of different societies can radically differ. That extends deeply into the fundamental assumptions about 'who' can make moral judgments and how they can be negotiated – as Kenneth Read (1955) noted of Gahuku-Gamu morality,

where the lack of personal individuality changes the types of moral relations in place from individual to distributive. Indeed, such claims have been foundational in relativism in anthropology. While many major tech firms will likely not consider Papua New Guinea to be a leading market for new technologies, understanding the 'playing field' for what is permissible with new technologies should be grounded in an understanding of such 'ontological' differences across relevant global markets' spectrum. Many commentators have noted meaningful, if less extreme differences, between individualistic Western societies that champion free choice and those of former Soviet nations or collectivist nations of East Asia (see Widlok 2004; Hefner 1998). Exploring these differences as a minimum are not only to avoid allegation of purely 'Western' notions of morality, but to identify the different ontologies and processes that govern moral decision-making in each. Given the resource constraints often placed on similar studies, we would suggest that such fundamental exploration can be guided by existing recent ethnographies of different cultures.

## 'Virtuous' Phenomena

Rather than focusing inquiry on the technology itself, in line with Widlok's (2004) framing of an anthropology of virtue, any study of 'grounded ethics' must explore precisely the 'virtuous' moments where moral dilemmas play out. In the case of understanding future hardware like AR wearables, for example, that might mean exploring moments of dilemma and negotiation regarding a range of social phenomena, including privacy, presence, agency, equality of information, and representation. Such topics could be explored through observations of moments where these 'virtues' are negotiated, like the sharing and visibility of space in a home or between neighbors (i.e. privacy), or how friends, families, and colleagues delineate expectations for presence in the context of smartphone ubiquity.

While software platforms and algorithmic products, as are common on social media, may initially appear further from observable 'virtuous' acts – when precisely does the harm on social media happen, for example? – we nevertheless see the same approach as being relevant for the development of these products (in addition to content moderation, rules, and more). To surface the moral facts that govern many of the challenges faced by social media in the age of populism – like the spread of misinformation, the incitement of hatred or violence, and more – one might, for example, study real-life negotiations of facticity, free speech, or content curation, in addition to experiences of encountering the 'other' or escalating/de-escalating conflict. Following the Durkheimian thread, the morals facts guiding their future boundaries lie less in law and formal debate, and more in observations of how these experiences unfold in fora on- and offline: in confrontational sub-Reddits, live protests, and conspiratorial YouTube channels, but also in mixed office cultures, parent-teacher conferences, and content selection at home. Understanding how these moments of virtue are negotiated points to the underlying standards and mechanics at work.

Yet this only covers *what* to look for, not *how*. While a foundational understanding of the moral systems affected may be explored through traditional ethnographic techniques, the challenge of understanding how and in which ways these moral systems could change – and ensuring these are anchored to the capabilities of a future technology – are more uncertain. So-called 'experimental philosophers' have explored other routes to testing the 'boundaries' of certain virtuous actions. In attempting to resolve the famed trolley problem – wherein an out-of-control train can kill one group or another based on the train conductors choice of

track – experimental philosophy researchers (Copp 2012) asked a large sample to answer the problem in a range of different permutations, e.g. that one group had the conductor's mother, that one group was older or overweight. By testing 'real' resolutions of the problem with a representative sample of the population, researchers were able to identify many of the contours and nuances which shape the resolution of the problem. While not 'real' in the sense of remaining a thought experiment – and more quantitative in nature than capturing the qualitative dilemmas – such experimental design points to additional ways to systematically explore the boundaries of moral intuitions, beyond how foundational dilemmas are experienced today.

In recent years, ReD has attempted to explore new ways to bring such experiments into real, lived experience. We have in recent ReD work explored designing 'social experiments,' combining approaches of experimental design from social psychology (see Isen and Levin 1972; Darley and Batson 1973) and the situated interactions of ethnomethodology, through methods like breaching (see Goffman 1971; Garfinkel [1967] 1991). This was tested most recently in a study on the related topic of the social acceptability of AR wearables. ReD researchers designed a 'trivia night' experiment to test the acceptability of the unequal distribution of information via AR glasses. Researchers used a live, planned trivia night as the setting, providing one pre-selected team with high-tech looking glasses and the answers to the night's trivia questions. Through gestures and other artefacts, participants 'simulated' receiving information through the glasses as they consistently answered correctly. While no complaints were lodged by other, uninformed teams during the trivia rounds, when the final results were counted – and prizes were to be awarded – the other teams reacted with uproar. Such dynamics revealed both the tension with social norms of calling out abusers – pointing to moral intuitions that people were less likely to expose – as well as the importance of the 'stakes' in creating a context in which those intuitions were broken.

In order to balance a baseline understanding of these 'virtuous phenomena,' and how they could change, we would advocate a balance of both ethnographic research into foundational instances of these dilemmas and similar veins of experiments to contextualize them within the technology, and better tease out the nuances, boundaries, and 'flexibility' of such elements – notably through three additional variables of *User, Context & Monetization.*

## Moral 'Users' of New Tech

Studies of new technology often focus on lead users using comparable technologies (i.e. for future AR products, heavy users of smartphone-based AR or wearables) or expected early adopters (i.e. frequent gadget buyers). While such users no doubt provide valuable insight into new applications for technologies or adoption drivers, their status precisely as 'lead' does little to evidence the more general moral facts and ethical decision-making that will eventually drive the reception of these technologies. One need only consider the initial excitement about Google's Glass from some groups to recognize that such a disconnect can be fatal. Instead, the target should be more representative 'mainstream' users – and, importantly, not only the 'users' themselves. Many moral quandaries – like the aforementioned 'trolley problem' – force a balancing of individual, group, and societal interests. We believe that a future ethics will need to understand how to balance the interests of purchasers, close and distant social groups, as well as unacquainted bystanders.

## Application Contexts

Coming in tandem with the need to explore the broad set of 'users' affected by the ethical decisions around a technology is an attention to the social and spatial distinctions which may also be inherent in deploying new technologies. As the trivia night example elucidates, bringing new technology into a space that is 1) shared, and 2) where information is viewed as 'valuable,' can drastically change the dynamics around what is and is not ethical. Were the same experiment repeated at home over a friendly game of Trivial Pursuit, the stakes might – although not always – be lower. Similarly, there are no shortage of examples – take watching pornography, for example – where behavior that is appropriate or acceptable changes widely from public spaces to the office to the home. Understanding, at a bare minimum, the difference between private (e.g. home alone), shared-private (e.g. friends' homes), shared-public (e.g. offices), and public spaces (e.g. malls, parks) will likely be relevant for many technologies.

## Monetization

While perhaps seeming more focused in scope, recent attention to the monetization of personal data and renewed criticism of the exploitative practices of companies – and technology companies, in particular (see Zuboff 2015, 2019) – suggests a particular sensitivity to pricing, data monetization, and related business model questions as altering the 'stakes' around a certain issue. This also accompanies a shift from attention to the user as 'purchaser' of services to technology companies aiming to deliver a continued 'experience' or 'relationship' with the user (see Amirebrahimi 2016) as a driver of revenue – implying a growing relationship between even broader engagement with a technology and the notion of its monetization. And all of this rests on top of long-recognized moral issues surrounding the role of money and broader forms of exchange in societies (see Parry and Bloch 1989). As a result, different strategies of monetization have become intertwined in what counts as ethical action. A future ethics will need to understand how 'financial stakes' of a buying or even engaging with a product impacts its moral role in society.

Taken together, we see these variables as a framework for identifying the key objects of study necessary in defining a 'grounded ethics' for a given technology – as well as the broader toolbox of methods needed to discover these ethics in practice.

## CONCLUSION: TOWARDS A PRACTICE OF 'GROUNDED ETHICS'

In this paper, we have aimed to address the much-discussed challenge of defining an ethics for developing and deploying new technologies and technology products globally – by shifting *where* such an ethics should come from. We have argued that the classical, *normative* approach to developing ethical frameworks – now guiding much of the approaches of major technology firms and related practitioners – does not sufficiently solve this problem, given that it leads to incoherent and inconsistent responses to the same problem, remains too open to interpretation in practice, and lacks the nuance necessary to guide practitioners as they make decisions. Rather, we have argued that a different epistemological approach – that of *discovery* – is needed in order to create a reliable system of ethics. Building of the growing field of the anthropology of ethics, we have located that discovery in the moral facts of

societies, but especially in the individual dilemmas and moral conflicts that elucidate the processes, systems, and practices by which ethics are developed – and what these systems suggest about the state of moral permissibility and its future flexibility & evolution. Finally, based on both theoretical and empirical examples, we have tried to synthesize the approach for a 'grounded ethics' into a framework to guide the research design of future explorations, notably: (a) the cultural foundations of the targeted societies, (b) the 'virtuous' phenomena likely affected by the new technology, (c) the ethical interests of different social groups, and who is the moral 'user' in each case, (d) contexts relevant to *where* ethics may be applied, e.g. physical sites, varied social groups and (e) moral notions around monetization. While this framework remains incompletely tested in full, in this paper's role as a 'catalyst' for the EPIC community and wider practitioners, we envision a future vein for research and praxis to activate this framework in order to refine it and better explore how to integrate it into contemporary technology practice.

Let us stop for a moment to explore that last word – practice. While we have discussed the practical challenges of past ethical approaches from major technology companies, we have yet to discuss what 'practice' could look like in a 'grounded ethics.' This inevitably touches on the more often discussed question of ethics for anthropologists and other social scientists: that of our own role, practices, and positions relative to the people we study and represent to others. Given the challenges that Metcalf et al. (2019) raise surrounding the practice of ethics within technology firms, and Amirebrahimi's (2016) concerns about the 'flattening' of ethnographic research, there are significant practical challenges to a grounded ethics, most notably: How does the ethnographers' study of moral facts and ethical processes not become an ethical 'rubber-stamp' for technology products or projects? And how can the toolkit of 'grounded ethics' not become an over-simplification of complex moral negotiations?

These are, of course, complex questions worthy of extensive theoretical reflections, original research, and practical experimentation. As a starting point, we take inspiration from two of Laidlaw's reflections on the practice of an anthropology of ethics:

> Ethics, as self-formation, intrinsically includes a practice of inquiry, and presupposes … an initial disjunction or difference between the self and one's teacher or exemplar. (2014, 216)
>
> …
>
> [A]nthropological thought, in particular the exercise of the ethnographic imagination, can be a mode of reflective self-formation, a form of spiritual exercise, and since it necessarily involves not only ironic detachment tempering whatever degree of understanding 'from the inside' we are able to achieve, but also necessarily a certain suspension and detachment from one's own knowledge and standpoint, it is an intrinsically sceptical one. (2014, 224)

In line with the framing of the negotiation of ethics that we had outlined in this paper, we see in both of these instances of 'self-formation' – the one, engaging with the 'ethics' of someone studied, the other, engaging with the 'ethnographic imagination' – a powerful foundation for engaging executives, technologists, and designers into the complexities of the moral and ethical negotiations they face. In that light, the ethnographer's responsibility becomes to ensure that the recipient of a 'grounded ethics' also engages deeply with the experience of 'self-formation' that both the content and medium for communicating

intrinsically should enable. Put in its simplest terms, that boils down to a question of format: the production of a 'grounded ethics' should be communicated in ways that ensure deep engagement with the same moral negotiations that future users will face.

With ethnographic practitioners working for decades in technology, there are no doubt no shortage of immersive presentation, workshop, and communication formats that could support such a practice – many of which have surely been discussed at length within the EPIC community. For the needs of a 'grounded ethics,' as a starting point we would highlight one form of knowledge production which, we believe, will be well-suited to the function of 'self-formation' while engaging with ethics: *the trade-off*.

Rather than representing moral and ethical findings as static facts – thereby reducing them to binary guidance to be followed or rejected – describing moral and ethical systems as a set of trade-offs has the advantage of immersing decision-makers in the same balancing and weighing of virtues and moral costs which informants themselves are likely to face. In practice, this amounts to representing 'grounded ethics' in terms of the lived experience of:

- *Competing Virtues:* The virtues the members of society must balance in a given dilemma, how much they weigh or 'pull', and the underlying factors shaping their relevance and weight

- *Costs:* The moral 'costs' that individuals in a society would experience as a result of one decision or another – pointing also to the costs that a company would incur in the same decision

- *Processes:* The dominant logics, negotiation processes, and actors considered and/or involved in engaging with such decisions

In representing findings as a nexus of these virtues, costs, and processes, a 'grounded ethics' can thus represent not only a binary indication of ethical-unethical, but come closer to forcing users to engage more deeply with the dilemmas faced from the perspective of the people likely to be impacted in the future. That such a 'grounded ethics' is not a static snapshot thus also increases its applicability long-term: by including the underlying factors shaping how society engages with morality, the framework can be adjusted to account for evolutions in society. This is only more relevant when one considers the often circuitous and winding long-term path that guides the development of major technologies, in terms of business model, customer, and applications or use cases: the balancing of trade-offs can equally shift to match the changing realities within an innovation process. There is good reason to believe that, as such, a 'grounded ethics' can become an integrated tool in the innovation process – just as ethnography has become for problem or opportunity definition.

This represents only an initial foray into imagining how a 'grounded ethics' could look, and how it might, in practice, resolve some of the challenges faced by current approaches to ethics in the development of technologies. We challenge practitioners – from the closer world of ethnographers and design researchers, but also, from further afield, technologists and technology executives – to seek out and discover the ethics that will drive their future decisions. Considering the scale of societal change that technologies promise and technology companies aspire to, a more thoughtful route to defining the direction for that societal change remains *out there*, waiting.

**Ian Dull** is a Senior Manager at ReD Associates, and a lead in ReD's technology and mobility practices. He focuses on long-term industry changes driven by technological and societal shifts, and is interested in new tools to help companies decide what they *should* do, beyond economics. Ian holds an M.Phil in Archaeology from the University of Cambridge. idu@redassociates.com

**Fani Ntavelou Baum** is a Senior Consultant at ReD Associates, where since 2018, she has advised leaders in finance, FMCGs, medtech, pharmaceuticals and nonprofits on how to ground solutions for commercial problems in human insights. Fani holds an M.St. in Philosophy from the University of Oxford. fnb@redassociates.com

**Thomas Hughes** is a Senior Consultant at ReD Associates. Thomas has helped world-leading Life Science companies improve their medical technologies by bringing them closer to the patient experience. Thomas is a medical anthropologist by training and holds a PhD in Anthropology from the University of Copenhagen. thu@redassociates.com

## NOTES

1. In a recent survey of UK tech workers, researchers found that 28% had seen decisions made about a technology that they believed would have a negative effect upon people or society. Among them, 20% went on to leave their companies as a result (Miller 2019).

2. Durkheim's focus on sanctions has given rise to a determinist view of ethics for some commentators, creating a norm-driven view of ethics. The argument goes that, if our ethical obligations are defined by our social duties, then there is no room for individual input and interpretation, the argument goes. Therefore, ethical facts are demoted to the status of norms *simpliciter*. James Laidlaw claims that ethics has been largely ignored by anthropologists with few exceptions due to the influence of Durkheim's deterministic vision of the moral fact (Laidlaw 2014). Laidlaw argues that until the early 2000s, most studies on morality and ethics in anthropology adopted more or less explicitly the so-called Durkheimian paradigm: ethnographic work consisted in the elucidation of a set of norms and values for a given group or society.

3. It is useful to understand social systems surrounding moral facts: they offer a stable reference point of acceptable norms and values, as well as an informed take of the context that has shaped these moral frameworks. This is indispensable knowledge for any understanding of moral facts.

## REFERENCES CITED

Amirebrahimi, Shaheen. 2016. "The Rise of the User and the Fall of People: Ethnographic Cooptation and a New Language of Globalization." In *2016 Ethnographic Praxis in Industry Conference Proceedings,* 71-103. https://www.epicpeople.org

Anderson, Ken, Maria Bezaitis, Carl Disalvo & Susan Faulkner. 2019. "A.I. Among Us: Agency in a World of Cameras and Recognition Systems." In *2019 Ethnographic Praxis in Industry Conference Proceedings,* 38-64.

Copp, David. 2012. "Experiments, Intuitions, and Methodology in Moral and Political Theory." In *Oxford Studies in Metaethics: Volume 7*, edited by Russ Shafer-Landau, 1-36. Oxford: Oxford University Press.

Darley, John M., and C. Daniel Batson. 1973. "'From Jerusalem to Jericho': A Study of Situational and Dispositional Variables in Helping Behavior." *Journal of Personality and Social Psychology* 27(1): 100–108.

D'Ignazio, Catherine, and Lauren Klein. 2019. *Data Feminism.* MIT Open Press.

Durkheim, Émile. [1924] 2010. *Sociology and Philosophy.* New York: Routledge.

Durkheim, Émile. 1975. *Textes: II. Religion, Morale, Anomie.* Paris: Editions de Minuit.

Eubanks, Virginia. 2018. *Automating Inequality: How High-Tech Tools Profile Police & Punish the Poor.* New York: St. Martin's Press.

Fassin, Didier. 2006. "The End of Ethnography as Collateral Damage of Ethical Regulation?" *American Ethnologist* 33 (4): 522-524.

Fassin, Didier. 2014. "The Ethical Turn in Anthropology: Promises and Uncertainties." *HAU: Journal of Ethnographic Theory* 4 (1): 429–435.

Fiesler, Casey, Natalie Garret & Nathan Beard. 2020. "What Do We Teach When We Teach Ethics?: A Syllabi Analysis." In *The 51st ACM Technical Symposium on Computer Science Education (SIGSCE '20), March 11-14, 2020, Portland, OR, USA.* New York: ACM.

Frenkel, Sheera, Davey Alba & Raymond Zhong. 2020. "Surge of Virus Misinformation Stumps Facebook and Twitter." *The New York Times,* 8 March 2020; updated 1 June 2020. https://www.nytimes.com/2020/03/08/technology/coronavirus-misinformation-social-media.html. Accessed 7 August 2020.

Future of Life. 2015. "Autonomous Weapons: An Open Letter from AI & Robotics Researchers." *Future of Life Institute* website. 28 July 2015. https://futureoflife.org/open-letter-autonomous-weapons/ Accessed 7 August 2020.

Garfinkel, Harold. [1967] 1991. *Studies in Ethnomethodology.* Cambridge: Polity Press.

Goffman, Erving. 1971. *Relations in Public: Microstudies of the Public Order.* New York: Basic Books.

Hefner, Robert W., ed. 1998. *Society and Morality in the New Asian Capitalisms.* New York: Routledge.

Isaac, Mike, and Sydney Ember. 2016. "Live Footage of Shootings Forces Facebook to Confront New Role." *The New York Times,* 8 July 2016. https://www.nytimes.com/2016/07/09/technology/facebook-dallas-live-video-breaking-news.html. Accessed 6 August 2020.

Isaac, Mike. 2020. "Facebook Moves to Limit Election Chaos in November." *The New York Times,* 3 September 2020, updated 22 September 2020. https://www.nytimes.com/2020/09/03/technology/facebook-election-chaos-november.html. Accessed 25 September 2020.

Isen, Alice M., and Paula F. Levin. 1972. "Effect of Feeling Good on Helping: Cookies and Kindness." *Journal of Personality and Social Psychology* 21(3): 384–388.

Jobin, Anna, Marcello Ienca & Effy Vayena. 2019. "The Global Landscape of AI Ethics Guidelines." *Nature Machine Intelligence* 1: 389-399.

Laidlaw, James. 2002. "For an Anthropology of Ethics and Freedom." *Journal of the Royal Anthropological Institute* 8 (2): 311–32.

Laidlaw, James. 2014. *The Subject of Virtue: An Anthropology of Ethics and Freedom.* Cambridge: Cambridge University Press.

Laidlaw, James. 2017. "Ethics / Morality." In *The Cambridge Encyclopedia of Anthropology*, edited by F. Stein, S. Lazar, M. Candea, H. Diemberger, J. Robbins, A. Sanchez & R. Stasch. https://www.anthroencyclopedia.com/entry/ethics-morality. Accessed 26 September 2020.

Lee, Nicol T., Paul Resnick & Genie Barton. 2019. "Algorithmic Bias Detection and Mitigation: Best Practices and Policies to Reduce Consumer Harms." Washington: The Brookings Institute. https://www.brookings.edu/research/algorithmic-bias-detection-and-mitigation-best-practices-and-policies-to-reduce-consumer-harms/. Accessed 6 August 2020.

Metcalf, Jacob, Emanuel Moss & Danah Boyd. 2019. "Owning Ethics: Corporate Logics, Silicon Valley and the Institutionalization of Ethics." *Social Research: An International Quarterly* 82 (2): 449-476.

Moss, Emanuel, and Metcalf, Jacob. 2019. "The Ethical Dilemma at the Heart of Big Tech Companies." *Harvard Business Review*, November 14, 2019. https://hbr.org/2019/11/the-ethical-dilemma-at-the-heart-of-big-tech-companies. Accessed September 17, 2020.

Moss, Emanuel, and Schüür, Friederike. 2019. "Tutorial: Ethics in Data-Driven Industries." Tutorial conducted at EPIC2019 in Providence, Rhode Island.

Parry, Jonathan, and Maurice Bloch, eds. [1989] 1996. *Money and the Morality of Exchange.* Cambridge: Cambridge University Press.

Read, Kenneth E. 1955. "Morality and the concept of the person among the Gahuku-Gama." *Oceania* 25 (4): 233-282.

Read, Max. 2017. "Does Even Mark Zuckerberg Know What Facebook Is?" *New York Magazine – Intelligencer,* 2 October 2017. https://nymag.com/intelligencer/2017/10/does-even-mark-zuckerberg-know-what-facebook-is.html. Accessed 25 September 2020.

Sample, Ian. 2018. "Thousands of leading AI researchers sign pledge against killer robots." *The Guardian,* 18 July 2018. https://www.theguardian.com/science/2018/jul/18/thousands-of-scientists-pledge-not-to-help-build-killer-ai-robots. Accessed 7 August 2020.

Simms, Ronald R. 1991. "The Institutionalization of Organizational Ethics." *Journal of Business Ethics* 10 (7): 493-506.

Tamminen, Sakari and Holmgren, Elisabet. 2016. In *2016 Ethnographic Praxis in Industry Conference Proceedings*, 154–174. https://www.epicpeople.org

VentureBeat. 2011. "Google's In-House Philosopher: Technologists Need a 'Moral Operating System.'" *VentureBeat*, 14 May 2011. https://venturebeat.com/2011/05/14/damon-horowitz-moral-operating-system/. Accessed 7 August 2020.

Vincent, James. 2019. "The Problem with AI Ethics." *The Verge*, 3 April 2019. https://www.theverge.com/2019/4/3/18293410/ai-artificial-intelligence-ethics-boards-charters-problem-big-tech. Accessed 7 August 2020.

Warzel, Charlie. 2020a. "What We Pretend to Know About the Coronavirus Could Kill Us." *The New York Times,* 3 April 2020. https://www.nytimes.com/2020/04/03/opinion/sunday/coronavirus-fake-news.html. Accessed 6 August 2020.

Warzel, Charlite. 2020b. "Mark Zuckerberg is the Most Powerful Unelected Man in America." *The New York Times,* 3 September 2020. https://www.nytimes.com/2020/09/03/opinion/facebook-zuckerberg-2020-election.html. Accessed 25 September 2020.

Wheeler, Melissa A., Melanie J. McGrath & Nick Haslam. 2019. "Twentieth century morality: The rise and fall of moral concepts from 1900 to 2007." *PloS ONE* 14 (2), e0212267.

Widlok, Thomas. 2004. "Sharing by Default? Outline of an Anthropology of Virtue." *Anthropological Theory* 4 (1): 53–70.

Zuboff, Shoshana. 2015. "Big Other: Surveillance Capitalism and the Prospects of an Information Civilization." *Journal of Information Technology* 30: 75-89.

Zuboff, Shoshana. 2019. *The Age of Surveillance Capitalism*: *The Fight for a Human Future at the New Frontier of Power*. New York: Public Affairs.

**THEMATIC SESSION**

# Scaling Ethnography

In this session, we consider how ethnography takes root and grows in organizations. We'll investigate the many models, meanings, and uses of ethnography developed and deployed, addressing their successes, failures, and ethical implications. This session focuses on the struggles we experience as we attempt to scale our practice in our organizations and industries.

Session Curators: Bec Purser, Heli Rantavuo, Frank Romagosa

# Enacting Scales

## Reflections from an Anthropologist Working in Asia's Ad World

TIFFANY TIVASURADEJ, *Ogilvy*

*The conference theme for EPIC2020 is all about scale. For many, scale will probably evoke images of sizing up, moving forward, getting better. But does scale carry the same meaning in all contexts? Could scaling back be the key to enacting scales successfully? And is it possible to enact scales when ethnography and the broader topic of anthropology are unheard among those around you? Reflecting on my own experience working in Thailand and China and my encounters with other design and business anthropologists working in Asia, I share an honest career narrative about enacting scales. My PechaKucha speaks truthfully about the struggles in applying ethnography, and inspires with learnings on how anthropologists can adapt the broader practice of anthropology and find ways to continue contributing to organisations across societies in Asia.*



'Inside Russian Doll' by Katy Brennan

**Tiffany Tivasuradej** is a Consultant at Ogilvy in Hong Kong. She holds a degree in biological anthropology from Durham University (UK). She consults for some of the world's most influential brands where she specialises in developing human-first strategy to help multinational clients tackle challenges around branding, marketing, and customer experience in Asia.
tiffany.tivasuradej@gmail.com
LinkedIn: www.linkedin.com/in/ycttivasuradej

# Scaling Experience Measurement

## Capturing and Quantifying User Experiences across the Real Estate Journey

REBECCA J HAZEN, PhD, *Zillow Group*
GENNY MANGUM, Zillow *Group*
TOM SOUHLAS, *Zillow Group*

*Zillow is undergoing a major evolution, transitioning from serving as the world's largest digital marketplace for real estate advertising into an end-to-end platform to support customers across the phases of buying, selling, and renting homes. As Zillow expands into more transactional spaces, the company has recognized the need to develop a clear and actionable understanding of users and their experiences as they interact with our products and services. To address this need, we set out to establish an Experience Measurement program to provide organization-wide visibility into how well our Zillow experiences meet users' needs and expectations as they progress through their real estate journey. This program will enable teams to gain insights at the intersection of attitudinal and behavioral experience data and lead us to our end goal of empowering informed decision-making across all levels of the experience and organization.*

*In this paper, we provide an overview of our approach to establishing an Experience Measurement program at Zillow. We focus on a small subset of Experience Outcomes (XOs) in an initial feasibility study to develop a program that would scale and drive impact in assessment and decision making across lines of business. Finally, we share lessons learned throughout the process of developing and validating the framework and discuss the impact of this work on current and future organizational outcomes.*

*Keywords: user experience measurement, experience outcomes, scaling experience measurement.*

## INTRODUCTION

Zillow is a real estate platform that provides customers with access to a broad range of real estate information and services for buying, selling, and renting homes. In recent years, Zillow Group has launched increased services to support customers through the more transactional elements of the real estate journey including rental applications and payments (Zillow Rentals), home loans and refinance (Zillow Home Loans), title and escrow services (Zillow Closing Services), and even buying and selling homes directly to and from Zillow Offers (figure 1).

Figure 1. Zillow Group Services (select)

## Motivation

Our interest in creating an Experience Measurement program grew out of a deep commitment to our customers. As a company, we strive to provide high-quality user experiences, and empower our customers as they navigate the transactional journey to unlock life's next chapter. This dedication is modeled through several of our core values ('Customers are our North Star') as well as in one of our company-wide Objectives and Key Results (OKRs) ('Deliver an integrated, joyful customer experience') (Workfront 2020). As we take on new and greater challenges and provide more complex features and services, we must hold ourselves accountable to our customers and deliver high-quality experiences. To do so, we must first define what makes for a good experience, and then measure the quality of those experiences in a comprehensive and actionable manner. As a result, we set out to create an Experience Measurement program that would shape how we think about customer problems and enable us to use measurement constructs to drive high-quality design and decision-making activities.

## Related Work

The value of understanding and measuring user experiences, and the importance of user-centered metrics, has been an active area of investigation in recent years. As a result, research and product teams around the world have developed and disseminated new frameworks and approaches to quantifying these elements. One such example is the HEART (Happiness, Engagement, Adoption, Retention, and Task success) framework developed at Google (Rodden 2010), which was designed as a means to leverage user-centered attitudinal and behavioral metrics across five categories to measure progress toward business objectives. This framework took the approach of developing an actionable understanding of how products and services are performing for the user through a synthesis of user-centered attitudinal and behavioral metrics and doing so at scale. Another example is the development of a measurement framework centered around Customer Experience Outcomes that was recently demonstrated at Amazon Prime Video (Morris and Gati 2018). In this paper, Customer Experience Outcomes (CXOs) were formulated based on an understanding of the jobs that customers were doing, developed through a series of in-depth qualitative and ethnographic studies. The authors then took a survey-based approach, coupled with behavioral analysis, to measure each of the prioritized CXOs at scale.

Scaling Experience Measurement – Hazen et al.

After examining our own user needs and business objectives, we decided to extend and apply similar concepts at Zillow in creating an experience measurement program centered around experience outcomes. For us, the program's success hinged on our ability to capture and objectively assess the quality of experiences as they occur across digital, physical, and human-based touchpoints. This meant identifying what was important and made for a good experience, as well as defining metrics and developing measurement plans to reflect the nature of these experiences. However, the challenge was how to make this a reality in a manner that would scale and keep up with the fast-paced and complex nature of our business.

## Program Objectives

Throughout this work, our objective was to develop a program to measure the quality of our user experiences and to empower stakeholders to use this information to drive high-quality design and decision-making activities. We aimed to provide stakeholders with a method of understanding how we are doing within our user experiences and a scaffold for digging deeper into the data surrounding these experiences to contextualize the results. We strived to support these stakeholders in evaluating the impact of design decisions, incorporating these framework elements into a range of evaluative studies from simple usability studies to large A/B tests. And finally, we wanted this framework to support prioritization efforts within product spaces and across lines of business.

This paper describes our process for developing the Experience Measurement program, and then walk through our experiences in an initial feasibility study. We share how the innate complexities of the activities and services offered often required us to rely on different qualitative and ethnographic techniques to investigate experiences and how we worked to achieve scale across diverse domain areas. Next, we discuss lessons learned as we refined our overall process, focusing on depth, efficiency, and ability to scale across the different experience areas at Zillow. We also talk about the value and role of ethnography to inform design and business decisions at Zillow. Finally, we share scenarios surrounding how this framework has been used to date at Zillow in order to (1) empower others to conduct consistent and high-quality measurement research, (2) inform the contextual underpinnings of dozens of attitudinal and behavioral metrics across the business, and (3) contribute to an expanding framework for how organization-wide prioritization will be conducted in the future at Zillow.

## APPROACH

### Experience Outcomes

The goal of this program was to develop a scalable framework for objectively measuring the quality of experiences across the end-to-end customer journey. The foundations of this Experience Measurement program are grounded in a partner construct known as Experience Outcomes (XO). Drawing inspiration from the jobs-to-be-done framework (Morris and Gati 2018, Christensen 2003, Christensen 2016), the basic unit of XOs are essentially the "jobs" that customers are attempting to accomplish and are employing our products and services to do. XOs are constructed through the extensive synthesis of research insights drawn from

interviews, observations, and other qualitative and quantitative user research conducted over time.

At Zillow, XOs are intentionally solution-agnostic and are designed to reflect the different personas or customer types who interact with our products and services. Each XO is accompanied by a detailed summary that describes 'what does success look like for this XO,' and 'what makes for a good experience when interacting with elements that support this XO.' We also include a description of how we support these elements within our current user experiences. This information not only helps to further contextualize the XO but also establishes the foundation for what will be measured in the Experience Measurement program.

The Zillow XO framework was developed by user experience research and design teams and was deployed across the company approximately nine months in advance of diving into developing and validating the Experience Measurement program. By that point, XOs had become a common currency of sorts, helping to provide context and create a shared understanding of the problems we were aiming to address through our solutions. With XOs now widely accepted and thus the foundation built, we decided that it was time to take the next step and fully operationalize the XOs into more actionable concepts that could ultimately be measured.

## 1. Developing a Framework

With these objectives in mind, we set out to create a measurement program that would transform our existing XOs into actionable concepts and operationalized measurement strategies. The result was an Experience Measurement framework (see figure 2).
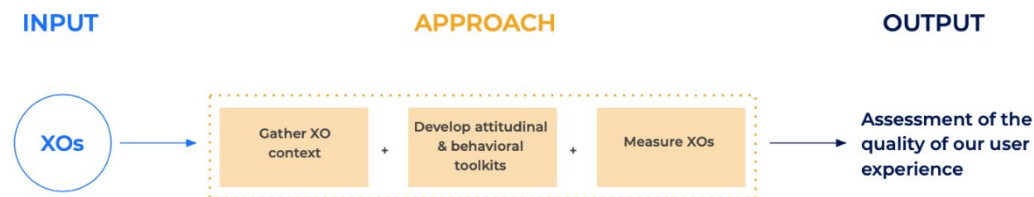


Figure 2. Experience Measurement Program Framework.

**1.1 Defining Key XOs:** The first element of the Experience Measurement framework consists of defining the prioritized or key XOs. At Zillow, we have over 150 different XOs representing the jobs being conducted by the various personas across the phases of the end-to-end real estate journey. We decided to focus on a prioritized set as operationalizing measurement plans for all XOs at once would be logistically impossible. Prioritization efforts involved a cross-team synthesis of research findings and insights. We then selected a prioritized set of XOs based on their importance to the customer and the opportunity space for the business.

**1.2 Defining Success Criteria and Mapping to Zillow Experiences:** In our framework, XOs are accompanied by success criteria: a set of high-level concepts that further define what makes for a good experience for a given XO. These criteria not only provide important context to those working in the space, but they also define the critical elements of success that we will measure at a more granular level later in the program. Once we define our success criteria, we then formally map out how and where these elements occur within our current user experiences, both on and off the Zillow website. This process helps to validate the criteria and lays the groundwork for future toolkit development, as decisions around how, where, and when to capture user data are considered.

**1.3 Developing Attitudinal Toolkits:** Early in developing this program, we determined that a comprehensive understanding of the user experience requires a synthesis of both attitudinal and behavioral inputs. As a result, we set out to create a set of attitudinal and behavioral toolkits to create structure and empower our teams with the tools required to develop well-constructed measurement plans. Attitudinal toolkits are designed to define approaches to capturing attitudinal data in support of our overall understanding and assessment of the effectiveness of our experiences and solutions. These toolkits consist of survey questions, measuring the high-level success criteria and XO satisfaction, and targeting strategies outlining where and when in the user experience these questions should be deployed to customers.

**1.4 Developing Behavioral Toolkits:** We also developed behavioral toolkits to guide the types of behavioral data to be collected and analyzed in order to support our overall understanding. The behavioral toolkits were developed in partnership with data science and analytics partners and were aimed at understanding and analyzing behavioral metrics and components closely associated with the success elements of the experience.

**1.5 Measuring on Experiences:** The final element of the program involved capturing data to drive analysis. In addition to the goals outlined in the program objectives, the outcomes of these analyses are intended to inform the prioritization of key XOs, reinitiating the framework process.

## 2. Validating the Framework

Our next step in this work was to conduct a small feasibility study to validate the framework, and generate lightweight, scalable protocols to facilitate framework adoption across the product spaces and experience areas at Zillow (figure 3).
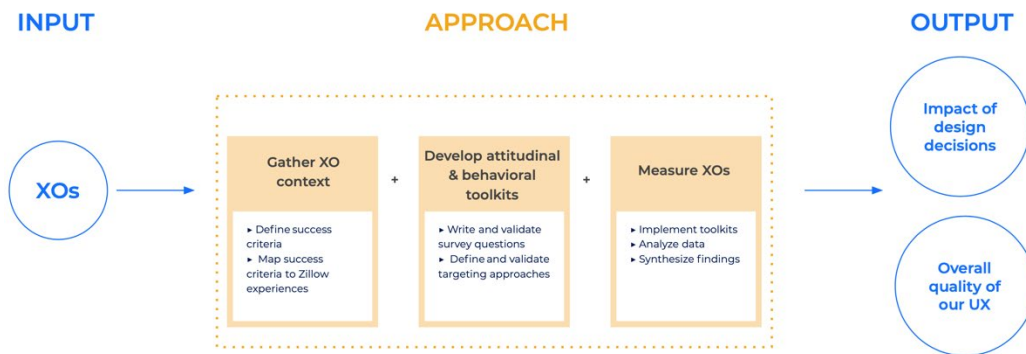
Figure 3. Experience Measurement Framework, expanded.

**2.1 Defining Key XOs:** We selected a small sample of XOs to focus on for the initial feasibility study. Unlike our typical prioritization process, we selected these XOs based on the anticipated value that they would provide toward our efforts to validate the framework and scale the supporting protocols. We chose certain XOs because we felt that they would generally be representative of other prioritized XOs based on the type and frequency of tasks involved. Others were selected based on potential challenges and opportunities that we anticipated, such as those involving smaller populations or doing more complex activities. We acknowledged that many of the activities associated with each of the initial XOs were taking place as on-site activities; we are currently in the process of planning a separate feasibility study examining the application of this framework on how we measure experiences that are taking place beyond the Zillow website and in scenarios that involve both on- and off-site experiences.

For the purpose of this paper, we will focus on a single XO: 'Help renters find homes based on stated criteria.'

**2.2 Defining Success Criteria and Mapping to Zillow Experiences:** Although we had previously developed success criteria for the majority of the XOs on our map, we determined that a new approach to defining these criteria was necessary for us to fully operationalize each XO for the purpose of this program. We went through a process of re-establishing contextualized success criteria for each XO, building a shared understanding with design and product partners of what success looks like, and what makes for a good experience. This approach relied on a variety of qualitative techniques, including interviews and contextual inquiry, to understand and explore these questions and capture the necessary context surrounding how users approach the task.

For the XO 'help renters find homes based on stated criteria,' we conducted semi-structured interviews with relevant individuals, coupled with observations to learn about the activities they were doing. We spoke with individuals who used Zillow as their primary resource for these jobs, and those who either did not prefer or did not use Zillow to capture a range of input and experiences. Through these activities, we probed on what this XO meant to users, and how they go about the task(s), and what makes for a good experience in doing so. We also took the approach of examining a parallel example, asking participants to

talk through similar questions for a similar activity in an entirely different domain. We found that this type of activity helped explore tacit elements of the experience and surface elements that people may value, but not fully realize or readily verbalize. We then examined what, if anything from these explorations, also applied to the rental scenario, then reconciled findings.

After exploring these questions and synthesizing the resulting insights into contextualized success criteria, we next examined the manifestations of these elements within our own user experiences. This activity aimed to create a shared understanding of where, when, and how these elements take place to inform survey targeting decisions for our toolkits and provide context to design and product stakeholders when interpreting findings from the XO measurement activities.

This manifestation activity was driven by insights arising from observational data and was conducted in collaboration with stakeholders from design, product, marketing, and analytics teams. In the case of the XO 'help renters find homes based on stated criteria,' this was relatively straightforward as the behaviors comprising this XO were easy to observe through standard user sessions. We had other XOs, however, that proved to be much more complicated. For these XOs, we had to rely on non-traditional techniques to indirectly "observe" users as scenario-based sessions failed to capture the range and authenticity of natural user behaviors. In one case, we opted to leverage FullStory (FullStory, Atlanta, GA), a digital analytics experience platform previously implemented within several product areas at Zillow, as it allowed us to remotely observe reconstructed user sessions associated with the XO. Although not initially implemented with this use case in mind, we found the tool to be useful to facilitate these observations as we felt that it would provide unique access to behavioral data and reduce biases associated with observation and scenario-based task observation. FullStory provided access to a great amount of unique and unobtrusive observational data, however, we acknowledged that the data was limited to digital interactions and did not provide any insight into verbal or emotional displays that were occurring as the users interacted with the user experience. To address the limitation, we then triangulated insights gathered from these observations against findings from the interview studies to analyze these emotions and behaviors and form a complete picture of the user experience. This activity helped to solidify our understanding and mapping of these manifestations before further operationalizing the metrics.

**2.3 Developing Attitudinal Toolkits:** We next started developing our attitudinal toolkits to assess the quality of our experiences and the effectiveness of our solutions. This involved creating survey questions and targeting strategies that we could use to capture attitudinal data about these user experiences. In our framework, the type and number of questions included may depend on the measurement scenario and objective. In each case, we took the approach of measuring at the level of success criteria element and XO satisfaction.

One important component of these toolkits is the ability to directly capture attitudinal data in the context of the user experience that is being assessed. Recognizing that a "one-size-fits-all" approach to data collection would be unlikely to meet the needs of the broader organization, we invested a significant amount of time and effort into determining both how and when to best capture data. This approach was technically challenging, but we believed it would provide us with higher quality, and more relevant, actionable data. The core element of this understanding was at what point in the experience has the user had sufficient

exposure to be able to confidently answer these questions. Determining when, where, and how to deploy in-context attitudinal data collection relied heavily on insights coming out of the observational studies and interviews, and behavioral data extracted in partnership with our data science partners. We also had to keep in mind the importance of data quality and minimize the impact on the overall user experience when developing these strategies.

In the case of our XO 'help renters find homes based on stated criteria,' we decided to use a site intercept survey that would prompt users for responses to a set of attitudinal measurement questions directly in the context of performing the activities of interest. The corresponding targeting strategy outlined details including the page to display the survey on, the amount of time delay before launch, and variables such as the number of pages visited prior to survey eligibility. We also worked with our data science partners to determine sample size and study duration, based on existing behavioral data. We then ran a qualitative study to validate both the survey questions (validity, reliability, comprehension) and targeting strategies before launching data collection.

**2.4 Developing Behavioral Toolkits:** The process of developing behavioral toolkits involved examining existing behavioral metrics, taking into consideration the user experience lens. Along with our product and analytics partners, we examined existing metrics and worked to determine how well each captured the elements of success we had previously outlined. In the case that metrics were missing or ill-defined, we proposed new metrics. Because behavioral metrics are often used to evaluate and communicate success across the business, we felt it essential to examine and validate these metrics.

**2.5 Measuring on Experiences:** The final activity was measuring on the user experiences, according to the strategies we had developed. For this study, this meant launching site intercept surveys and capturing corresponding attitudinal and behavioral data. For the feasibility study, we captured over 5,000 complete responses across our desktop and mobile web platforms. We analyzed these attitudinal data in conjunction with basic behavioral metrics to develop a high-level assessment for the XO and then explored relationships to identify additional research opportunities to contribute to our overall understanding (research synthesis).

## LESSONS AND IMPACT

### Lessons Learned

Although we have had many successes throughout this work, we have also encountered several learning opportunities along the way. One lesson that surfaced early was the value of gaining buy-in and alignment with cross-disciplinary partners early in the process and continuing to foster that sense over time. We found that establishing initial buy-in through reflecting value and alignment with the things important to each of the stakeholder groups, and then building on that sense by including stakeholders in conversations and research activities throughout building toolkits provided benefit beyond expectation. Having participation from these groups along the way also made it easier for stakeholders to see potential applications, and to contribute to shaping the overall program and further informing outcomes. We also found that working from an existing framework (XOs) that

many people across the company were already familiar with helped gather early support/buy-in.

One aspect that we had to spend a relatively significant amount of time and energy on was finding the balance between quality and flexibility in creating repeatable research protocols to truly achieve scale. This program is a massive undertaking that requires contributions from dozens of researchers and product teams across domains of the business. Because we did not have the luxury of having everyone drop everything and focus on building out toolkits for a month or two, making these activities easy to adopt as a part of current research flows was important. This was initially a challenge, however, we found success through identifying opportunities to more naturally incorporate these elements within common research activities and were able to create and demonstrate variations on the research protocols, depending on factors like time and availability of resources. This made adoption much easier to envision, and also helped to make rapid progress within a more lightweight and streamlined process.

Similarly, because moving fast is important, it was also incredibly beneficial to leverage existing knowledge and tools whenever possible. This helped us to more quickly operationalize the XOs and gain access to insights that might otherwise be difficult to capture.

## Organizational and Business Impact

Though still relatively early in this work, we have already seen a significant amount of buy-in across the organization, including within Senior Leadership, and an impressive rate of adoption among research and product teams. Since the initial feasibility study, we have empowered researchers across our team to engage in the process of building out attitudinal toolkits for a set of nearly 40 prioritized XOs across several key lines of business, using consistent and scalable research protocols. We are also scaling behavioral toolkits for these same XO areas. This has allowed us to move quickly in introducing this program to a broad range of teams, promoting the value of user experience measurement and XOs across the organization. We have also seen teams sharing findings related to early measurement work; through this, are beginning to see evidence of how the XOs and measurement can be used on a larger scale to help inform planning and prioritization efforts among design, product, and business teams.

Another major outcome that we are seeing is the value of leveraging insights gathered throughout qualitative and ethnographic studies to encourage conversation and evaluation of existing metrics in place across the organization. This served to inform the contextual underpinnings of dozens of attitudinal and behavioral metrics relied upon by team across the entire business. Historically, teams have relied on these types of metrics business they are easy to communicate, and quickly and concisely convey meaning. By encouraging conversations and reframing these metrics around the concept of user experience and measurement on XOs elements, we are further reinforcing the value of user experience as a higher-level factor in our metrics and success as an organization.

## CONCLUSION

We started out in this work with a goal to develop an actionable understanding of how well we are meeting the needs and expectations of customers as they interact with our products and services across the real estate journey. As a business, and within individual product spaces, there is also a need to know what to prioritize or focus on next, often in the face of competing priorities. We felt strongly that grounding these decisions in the user experience aligned with our values and would enable more informed decision making. Through this work, we developed a scalable Experience Measurement program for defining survey strategies and driving high-quality contextual data collection across experiences in the real estate customer journey. In leading with a qualitative, ethnographically-informed approach, and leveraging data and technology solutions already in place, we were able to move fast to meet organizational needs without compromising on research depth or quality. This approach also served to build empathy and strengthen the cross-disciplinary partnerships that were essential to the success of our ability to measure these experiences. We believe that this framework will empower teams to quickly and effectively define and capture data to evaluate the success of product, design, and business decisions, a capability critical to our evolution as a company.

## AUTHORS AND CORRESPONDENCE

Correspondence or requests for additional information can be sent to the primary author (RH) via LinkedIn (LinkedIn.com/in/RebeccaHazen5349) or email rebecccahaz@zillowgroup.com.

## REFERENCES CITED

Christensen, Clayton M, Taddy Hall, Karen Dillon, and David S. Duncan. 2016. "Spotlight on consumer insight: Know Your Customers' "Jobs to Be done." Harvard Business Review.

Christensen, Clayton M and Michael E. Raynor. 2003. "The Innovator's Solution: Creating and Sustaining Successful Growth." Boston, MA: Harvard Business Review Publishing Corporation.

Morris, Lauren and Rebecca Gati. 2018. "Humanizing Quant and Scaling Qual to Drive Decision-Making." *Ethnographic Praxis in Industry Conference Proceedings*: 614–630.

Rodden, Kerry, Hilary Hutchinson, and Xin Fu. 2010. "Measuring the User Experience on a Large Scale: User-Centered Metrics for Web Applications." *Proceedings of CHI 2010*.

Workfront. 2020. "What are OKRs?" https://www.workfront.com/strategic-planning/goals/okr

PECHAKUCHA

# One Small Step for Ethnography, One Giant Leap for Banking and Insurance

JENNIFER ROTH, *USAA*
SARA KLUCKHOHN, *USAA*

*In this presentation we argue that in many regulated industries such as banking, finance, and insurance, a post qualitative vs. quantitative world is not yet a reality.  In such an environment, advanced analytics could be likened to being in its teenage years, while behavioral research is still in its infancy.  Big data primarily drives our metrics, but in such a highly digitized and individualized culture, we know that ethnography is the missing piece of the puzzle. This means that as social scientists we must be the loudest (and sometimes lone) voice calling to leverage employees who are trained in these skill sets and incorporate these methods into our work. Slow and steady wins the race and our wins look different when compared with companies that already have been convinced of the value and don't have to do as much work to incorporate them into existing analysis. We have found that becoming EPIC members has been a turning point for our own growth mindset.  Our industries are still primarily relationship-based, and for those analysts new to regulated industries, EPIC is the ideal community to help propel the legitimization of a permanent place for qualitative methods in our data-driven, highly regulated industry.*

*Keywords: quantitative methods, mixed methods, research design, finance, insurance, regulated.*

**Jennifer Roth, PhD,** is a Senior Business Strategy Analyst at USAA working in analytics related to property claims. Jennifer has an MA in anthropology and a PhD in applied demography; she is also currently working toward a Chartered Property and Casualty Underwriter designation. Jennifer has been a champion for the use of social science in corporate settings since starting her career and is often sought out by leaders for her assistance in research design. She is very grateful to have Sara as her partner in this journey!  Jennifer can be reached at Jennifer.roth@usaa.com.

**Sara Kluckhohn** is a Senior Decision Science Analyst at USAA, a major auto and home insurer, where she tracks and reports on customer experiences. Sara has an MS in sociology and is studying for the Certified Specialist in Predictive Analytics exam offered by the CAS Institutes. She and Jennifer have been colleagues and friends for 5 years and often partner together in cross functional efforts to support each other's research. Sara can be reached at sara.kluckhohn@usaa.com.

**THEMATIC SESSION**

# Transcending Time and Space

In this session we consider how ethnography and ethnographers journey across scales. We'll consider how ethnography (and ethnographers) can escape the present to explore futures, and examine different ways of dealing with distance.

Session Curators: Laura Cesafsky, Kat Ekberg, Evan Hanover, Scott Matter

# The City as Organization

## Ethnography for Alternative Futures

JORDAN SHADE, *International Business Machines Corporation, A Functional Democracy*
HAL WUERTZ, *International Business Machines Corporation, A Functional Democracy*

*In this case study we use ethnographic outcomes from the study of the employee population of IBM, to inform new experiences for improving civic engagement in the resident population of Austin, Texas. In doing so, we experiment with a technique in speculative ethnography that uses research insights from a variant population with a variant challenge for in-depth explorations of a possible future. We demonstrate, first, that while in speculative thinking big ideas can be imagined, transposing ethnography enables a richer exploration of possible futures, and thus, further depth in ideas. And second, that by combining speculative thinking with existing ethnography, researchers and design teams can unearth bold experiments and jump start a design process that drives quicker learnings and impact in new contexts.*

*Keywords: Culture Change, Speculative Design, Civic Engagement, Design Research, Anticipatory Anthropology*

## INTRODUCTION

The largest scale problems in the world are often challenges of collective behavior. How do we change behaviors that affect climate change? That perpetuate gender inequality or racial bias? That hinder productive civic dialogue and participation? In this case study, we consider culture change for the public good: improving participation in our democracy. More specifically, we tackle this challenge at the local level in our fair city of Austin, Texas, where, like in my other places in the United States, civic participation poorly reflects the diversity of the population.

The approach we use combines speculative design, ethnography and human-centered-design to propose radical solutions to Austin's civic engagement challenges. In order to accomplish this, we draw from many years of research informing cultural change at IBM, in order to advance a cultural change mission in Austin, Texas. In juxtaposing these two change programs, we're able to leverage both their similarities and differences to surface unexpected possibilities and interesting questions about what is given, what is ideal, what is likely, and what is possible. The result of this work is a very tangible set of outputs that help tackle the substantial challenge of how to change behaviors around civic engagement. In the process, we understand new ways to scale ethnographic insights about human behavior by re-applying them in new contexts, allowing us to address our world's most wicked problems with deeper understanding.

## PARALLEL PROJECTS IN CULTURE CHANGE AT IBM & IN AUSTIN

From 2014-2019 we worked as senior designers and researchers at IBM Design towards the explicit goal of creating a sustainable culture of design and design thinking, to help transform IBM into a more human and design-centered company, as a part of the Design Program Office (DPO).

To understand existing behaviors, how to best drive adoption of a new practice (Enterprise Design Thinking), and what conditions were needed to achieve both cultural and business impact, we worked through a multi-year longitudinal research study across the company. Methods applied include:

- Hundreds of hours of design studio ethnographic observations (14 studios studied globally)
- Multi-stage generative and evaluative interviews of product teams, consultants, sellers, executives and practitioners
- Surveys of team members who had been trained 30/60/90 days out of training
- Participatory design workshops with experts within the business
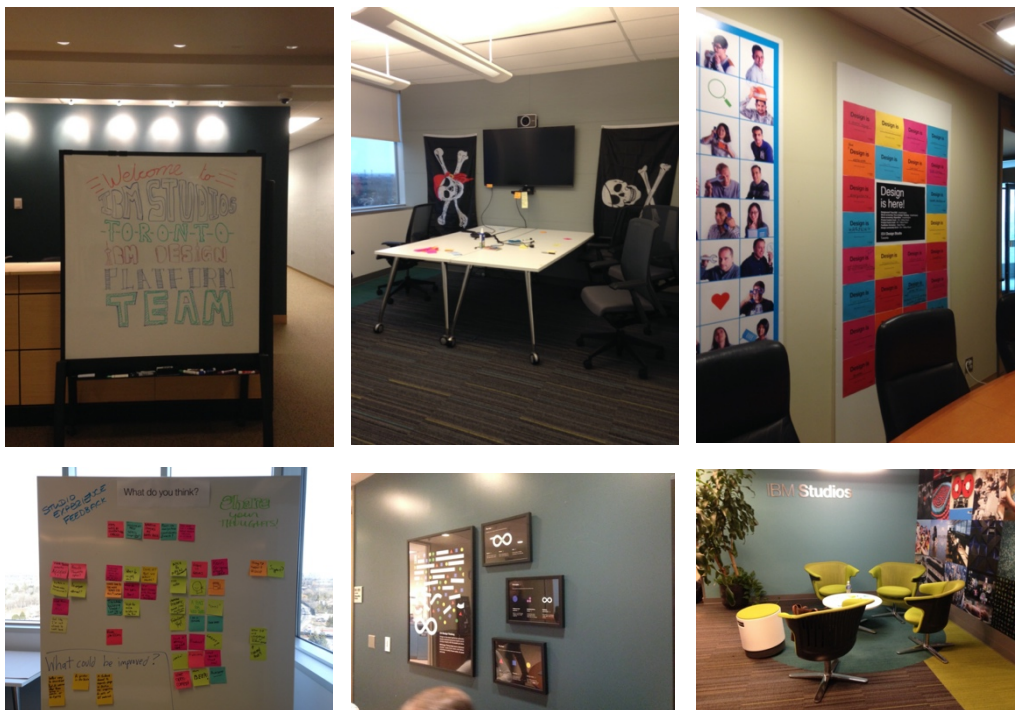


Figure 1: Photos from various IBM Studios research trips
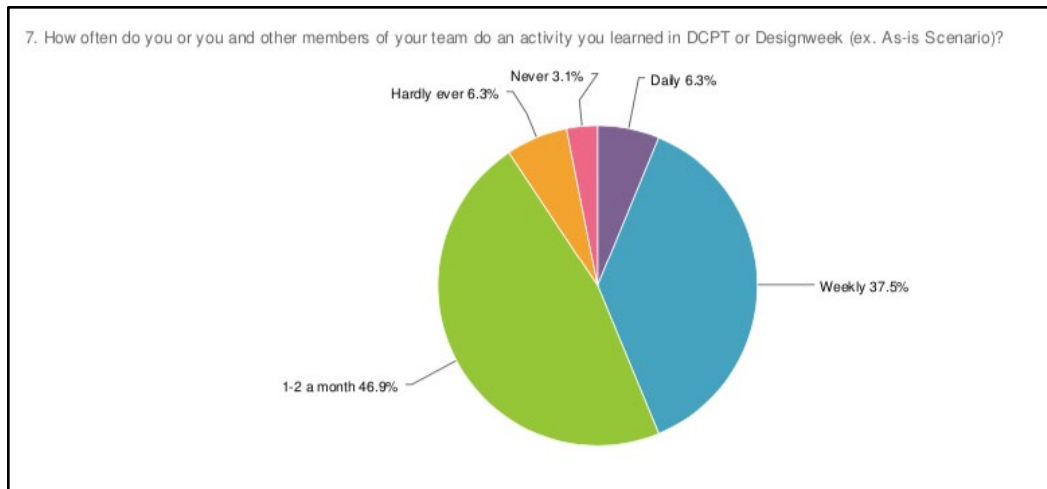
Figure 2: Example survey question and answers

## IBM Key Insight #1 : Ecosystems of Adoption

At IBM, we learned that adoption of Enterprise Design Thinking[1] was fueled by the presence of catalysts we internally dubbed "Magic People"—experts in the practice who could facilitate and encourage others, who were organically spread out around the business. But we saw they were drowning in the additional work of transforming the communities around them. They needed others, at various levels of expertise and authority to support their work.

These "Magic People" (and their needs) became the heart of the successful Enterprise Design Thinking ecosystem: EDT Coaches who could support and lead the Practitioners and Co-Creators around them, but be supported and protected in turn by Advocates and Leaders.

However, we also uncovered that we couldn't just "train the trainer" and mass produce these "Magic People," or even entry level practitioners. There was no time to educate every IBMer to the appropriate level with in-person learning—we needed to provide a journey and path for employees to move themselves from "level 0" to their appropriate place. Without such a path, it was too easy for motivated individuals and teams to stall and eventually become disillusioned at the total lack of change in their sub-communities. For the ecosystem to work, some would need to grow into advanced leadership levels, while many would become solid day-to-day  practitioners, but they all needed to see the way forward and get education at each stop along the way.

This journey became the connective tissue between the digital badges, the surrounding context that informed desired ratios between the badge levels, and what enablement and resources were needed as IBMers progressed through their personal journeys.
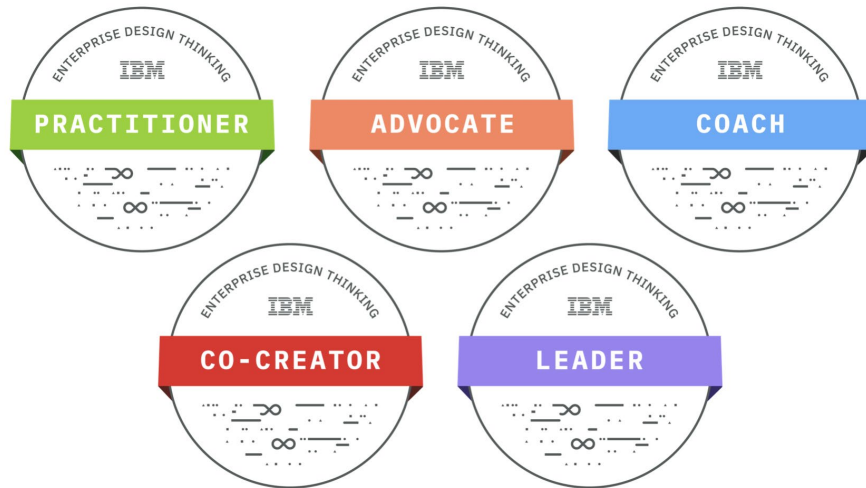
Figure 3: Enterprise Design Thinking badges system, designed in early 2017
(IBM.com/design/thinking)
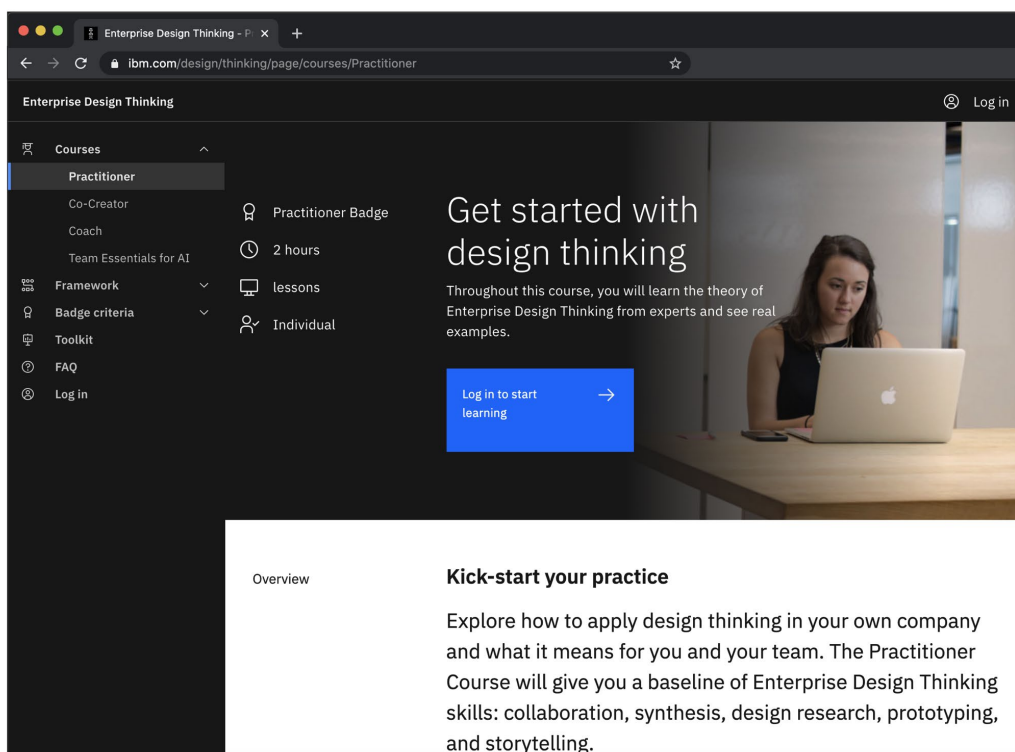


Figure 4: Enterprise Design Thinking online learning modules (IBM.com/design/thinking)

## IBM Key Insight #2 : Emergent Communities & Ownership

As we continued to observe the "Magic People" or EDT Coaches, we saw they needed to be connected together through a support system tied to a central community of practice where they could share experiences and bond. We also needed to help broadcast them to their local communities via what became a published set of EDT Chapters: decentralized hubs providing support and guidance leveling up, that were based on emergent communities spun up around known experts in various studios and geographies.

Often as existing experts, who were deft at nuanced customization of the practice within their specific business contexts, the EDT Coaches were both highly dependent on the centralized DPO for support, as well as resistant to its control and perceived oversimplification of technique. We found quickly that mandates don't work very well even (especially?) in corporate environments. Something the Coaches knew inherently was that people need to see the value in their reskilling for their everyday activities, and then they will voluntarily reskill. For this reason, we didn't use mandates very often to teach design thinking at IBM, but instead relied heavily on these sub-communities where personal relationships and local recognition pulled people into a transformational effort in a very intimate and human way.
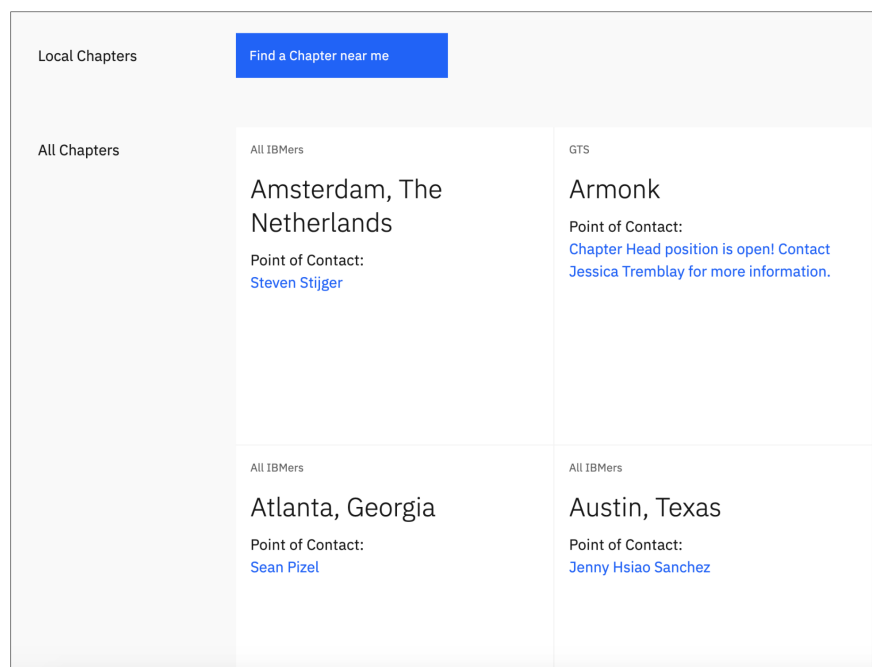


Figure 5: In 2020, there are 105 decentralized Chapters in the network directory.

## IBM Key Insight #3 : Studios & the Impact of Scaled Assets

Another component of our change program is "IBM Studios"—a global network of 50+ spaces designed to house designers and design thinking processes. Studios were

designed for collaboration and thus feature open-office style seating and large spaces to host internal and external co-creation activities. We found that Studios served a dual purpose, not only as a functional change mechanism but also as an important symbol of change to both internal and external audiences. This symbol was instrumental for internal morale as well as for clients who couldn't always see the day-to-day workings of IBM, but could visit a studio for a 1-hour briefing. For this reason, we took special care to recognize Studios as a symbol of the transformation, and so supported Studios in branding, creating visitor experiences and telling the unique story of each individual space.

The Studios became hubs for the use and distribution of assets produced by the DPO: educational materials, workshop decks, client references, books and toolkits, as well as the home for many of the EDT Chapters. We found repeatedly that these communities were hungry for reusable assets and so as a design team a key part of our work was identifying patterns of success, designing subsequent assets, and making them available to the global network. It was common for these branded assets to have a meme-like effect where they were used without meaningful knowledge of the underlying intention. We began to anticipate this outcome and considered this fact in how we designed and distributed assets.

## IBM: IMPACT AND RESULTS

We've seen massive success in this cultural transformation. IBM tops the list of educated design thinkers at a single company with over 125,000 employees skilled up in Enterprise Design Thinking, and 20,000 formally trained designers. We've seen an 11+ increase in IBM's NPS scores in 2019. Forrester reported faster, more efficient and higher quality outputs, documenting the impact of these new cultural processes[2]: product deployment on average 2x faster to market, a 300% return on investment and a 75% increase in efficiency.

What's more, years in, we see the practice grow organically on its own, through human connections, an adoption that takes on a life of its own, and no longer depends on a centralized initiative. For example, the Enterprise Design Thinking Slack channel is the single largest at the company, boasting over 15,000 active members, and daily discussions on a wide variety of topics between IBMers who don't typically know each other, or work together, but are deeply engaged in growing their individual, team and company practice.

### The City of Austin: Creating a Culture of Civic Engagement

While big business is our day job, after the election, beginning in 2016, we had started a passion project: a mission based organization called A Functional Democracy, with our partner Amy Stansbury. Realizing the limited way that Austinites were engaged at the critical local levels of government, where we can actually make more impact on policy, regulations, and budget, we originally looked at re-designing resources provided by the city—resources that struggled to communicate immediate ways to participate, or the system of local government—without using insider language and ancient interaction paradigms.

Over time we evolved towards the mission of creating a culture of local civic engagement and activism. Through this work we empower local residents to take actions to make their city more directly reflect their needs and values.

**PROCEDURES FOR AUSTIN CITY COUNCIL
MEETINGS, PUBLIC HEARINGS, AND APPEALS**

**ARTICLE 1. PROCEDURES FOR CITY COUNCIL MEETINGS.**

**Division 1. Regular Meetings.**

§ 1-1-1 COUNCIL FINDINGS.

The Council finds that:

(1) Useful public debate at its meetings is served by the fair and impartial administration of rules of order.

(2) Democracy is best served by participation of people in their government.

(3) Austin's exercise of representative democracy is well-served by the timely and substantial participation of its citizens.

(4) Minor, incidental breaches of decorum may occur in the exercise of participatory democracy.

(5) The best government is that which is done in the open.

§ 1-1-2 PROCEDURAL DIRECTIVES.

(A) These Procedures are directory only, and a violation of any of the provisions of this article does not affect the validity or legality of any Council action otherwise properly adopted.

(B) The Council may waive any of the requirements of these Procedures for a particular agenda item by voice vote without the need to amend the Procedures.

(C) The city clerk shall make copies of these Procedures available to the public.

§ 1-1-3 RULES OF ORDER.

(A) Except in the case of a conflict with these Procedures, the City Code ("Code"), the City Charter, or state law, Council meetings should be conducted under Robert's Rules of Order.

(B) Each person and Council Member attending a Council meeting should observe decorum. A person or Council Member should not speak out of turn, use disparaging or abusive language, or make threats of violence against any other person during a Council meeting.

(C) The presiding officer:

(1) should maintain order;

(2) should exercise the officer's authority impartially; and

(3) may shorten a person's speaking time or ban a person from speaking for the duration of a meeting only for a violation of decorum set out in this section.

Page 1 of 10
*Effective May 29, 2017. Revised June 22, 2017.*

Figure 6: A resource from the City of Austin website's City Council Information Center
(https://www.austintexas.gov/department/city-council/council/council_meeting_info_center.htm)

As we perused these two cultural transformation projects, the two worlds of IBM and Austin started to intermingle  for us as researchers seeking to understand populations and their participation in new practices.

Despite the obvious differences between these two populations, there was one glaring similarity: they both had cultural transformations under way. They both needed to undergo change to redefine their ways of working to be successful in the future. Towards this end, in both cases we were educating and empowering people to change the cultures around them for their own benefit. And thus, as researchers, we were looking for insight into how these populations learn, how new behaviors are adopted in groups, how new memes spread, and most importantly—how to create these behavioral changes at scale.

What was the relationship between these two change missions and how could we leverage our 5 years of ethnography at IBM for the cultural transformation of the city of Austin?

## "WHAT IF... " SCENARIOS

As part of our design process for creating educational tools for Austin, we started giving ourselves the freedom to directly apply what we'd learned about culture change at IBM to explore ideas for cultural change in the city. Doing so cracked open opportunities, and helped us to think more expansively about how large populations learn and how we could empower people to adopt new ideas.

We began brainstorming ideas for Austin starting with a few simple what-ifs inspired by our work at IBM:

- What if, in the same way that IBM hired 2000 designers to kick-start its transformation, we hired 2000 full time "Civic Catalysts" in Austin that would funnel citizen voices into the system?
- What if, like how IBM built studios to foster collaboration, we built civic studios in Austin, that activists could use to work together and connect the dots between their different missions?
- What if, like how IBM built a pathway for Enterprise Design Thinking, we built a framework for civic engagement levels in the city that was so widely adopted that we could use it to measure and track learning?

The ideas that arose from this process were interesting to us because they hit a unique sweet spot between possible and absurd—a perfect combination for futurist thinking. We decided to treat these resultant scenarios as opportunities for Speculative Design. Speculative Design is a term coined by professors Anthony Dunne and Fiona Raby, and popularized in their 2013 book, "Speculative Everything: Design, Dreaming and Social Dreaming." They describe Speculative Design as the following:

> "[Speculative design] thrives on imagination and aims to **open up new perspectives** on what are sometimes called wicked problems, to **create spaces for discussion and debate** about alternative ways of being, and to inspire and encourage people's imaginations to flow freely. **Design speculations can act as a catalyst for collectively redefining our relationship to reality."** (Dunne and Raby, 2013)

Rather than simply treat these "What-if" ideas as metaphors, as in a "Big Idea" brainstorm[3] where analogies are used as inspiration, we considered these "What-ifs" as real examples of how Austin might be in 10 years. These were possible futures that we wanted to explore in more depth.

Figure 7: In "What if" scenario 2 we questioned, "What if there were a framework for civic engagement levels for the city of Austin? And what if almost everyone in the city of Austin had a badge so that we could use this badging framework to recognize and track civic engagement?" (Source: "SXSW on 6th Street" by Ian Aberle via creativecommons.org and mediashift.org)

## APPLYING SPECULATIVE ETHNOGRAPHY

In order to examine these what-ifs more closely, and actually move towards prototyping in the real world, we apply a method we call Speculative Ethnography, wherein we hypothetically "observe" human behavior in a speculative future for one organization (City of Austin) based on what we know about human behavior in the past from a second organization (IBM). We define Speculative Ethnography in the following way:

> Speculative Ethnography is the informed study of how people could behave in future scenarios, based on what we know about human behavior in the present.

This approach of Speculative Ethnography is related to existing approaches that combine futures thinking and anthropological techniques including Ethnographic Futures Research (EFR), anticipatory anthropology and ethnofutures.[4][5] These related disciplines or concepts use evidence from today to think systematically about possible futures.(English-Lueck and Avery, 2020)

This approach also has similarities to the practice of Ethnographic Analogy, commonly used in anthropological study, in which observed anthropological data from comparative populations is used to create hypotheses about past human societies. (Currie, 2016) While in Ethnographic Analogy parallels are drawn to reconstruct past human societies, in the process described in this case study, a parallel is used to construct future possibilities.

Figure 8: Two artifacts from ethnographic design futures work. On the left, a participant creates a rapid prototype of 'personal experiential futures'. (Project design & photo: Maggie Greyson) On the right, Kelly Kornet creates an artifact based on research into the preferred future of an environmental activist. (Design & Photo: Kelly Kornet) (Source: Candy and Kornet, 2019)

To explore this Speculative Ethnography approach we look to an analogy from Star Trek, the television show set in the 23rd century. One of the key features aboard the Starship Enterprise is the Holodeck. Through this device, participants ask a computer to craft virtual reality scenarios based on previous knowledge so that participants can experience these scenarios in an immersive environment. While this device is mostly used for entertainment, the show uses it to explore philosophical questions, and even solve problems in the real world.

In this Speculative Ethnography exercise, it's helpful to imagine the what-if scenarios proposed as akin to the Holodeck. Using past ethnography as input, new scenarios can be created that rely on previous learnings. Via this method, ethnography is used twice in a recursive loop. First, to gather insights from a primary population, and then once a new "what-if" scenario is imagined, ethnography can be used to "observe" the secondary population. This method requires as input the in depth study of an existing population, and an analogous population.



Figure 9: The holodeck is a device from the television show Star Trek. With this device, participants may design and experience specific virtual reality environments. (Source: Screenshot via mediashift.org and screenshot via space.com and memory-alpha.fandom.com)

## An Example of Speculative Ethnography

We've used this technique to inspire several projects in A Functional Democracy. Here we walk through one end-to-end example of how we have used this method in our process from speculated future to prototype.

At IBM, one of the most powerful change management techniques we created based off of ethnographic research was a learning pathway and badging system. By studying teams we observed that there was a natural ecosystem of design thinking skills and roles, that when at play, accelerated the successful adoption of a user-centered approach.

Transposing this idea onto the City of Austin, we posited: What if there were a framework for civic engagement levels in the city of Austin? And what if almost everyone in the city of Austin had a badge so that we could use this badging framework to recognize and track civic engagement? We put ourselves into the Holodeck and pulled out hypothetical learnings about this scenario, as illustrated below.

## Hypothetical Observation #1

People are showing off their badges. They are proud of their status and make things like bumper stickers to represent their achievements. A counter-culture emerges that is anti-authority and vocally pushes back against the movement.



Figure 10: Illustration of opposing bumper stickers (Source: "Mur's bumper stickers and plates, Raleigh airport, Raleigh, North Carolina, USA.JPG" by gruntzooki and "Twins" by dave_stone via Flickr)

## Hypothetical Observation #2

The official badging system leads to the organic rise of online communities in sub-reddits, informal Slack communities, and on Nextdoor. Residents discuss politics in a new way using their new shared language, all focused on building community around local civic engagement, decentralizing civics education as a result.
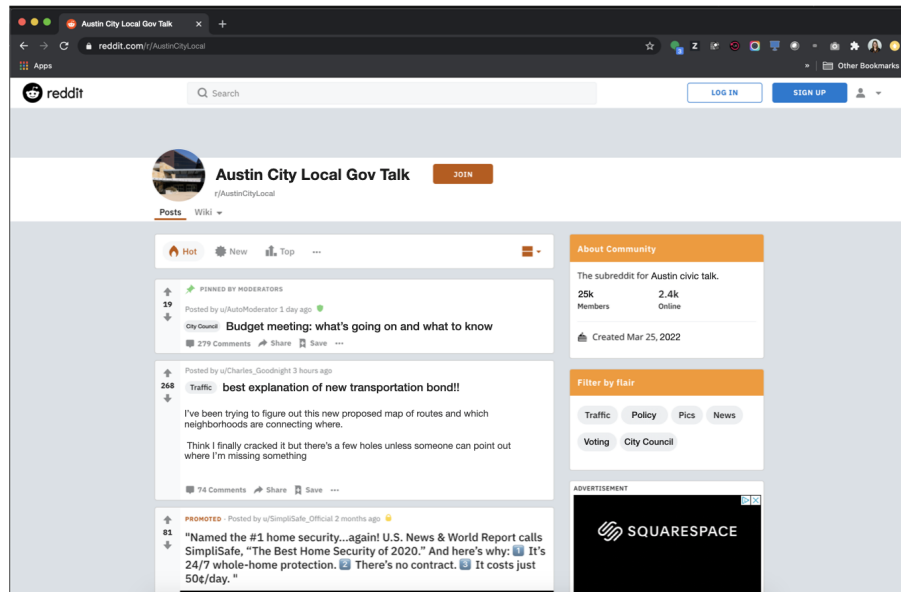


Figure 11: Illustration of local government subreddit

## Hypothetical Observation #3

Residents eat up scalable assets. People share the assets freely and use them often, both for their own education and for evangelism. But the assets are also fetishized. They have a meme-like effect; the symbols are more wide-spread than actual change.
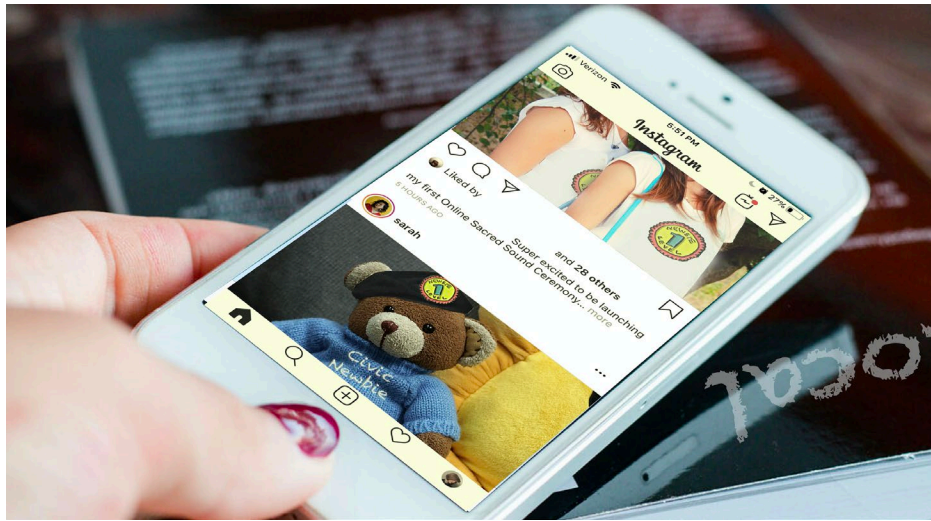
Figure 12: Illustration of the oversharing of the superficial symbols of the program (Source: "TouchID on Apple iPhone" by wuestenigel via Flickr)

## Holodeck Insights and New Questions

As we play out this scenario, we are able to "observe" and examine the ideas in play based on what we know from these ideas existing at IBM, and develop hypotheses for how ideas would or wouldn't work. We expect for example that some community figureheads (especially those who are experts in tangential disciplines such as advocacy, community organizing and activism) will view this new system with skepticism and want to align it to known frameworks.  We expect novices who have been seeking guidance in this area, but feel stunted in their own practice, will gravitate to this framework immediately. We expect that they will share it widely and start to experiment with its activities to fit their own needs. We also believe that the accompanying educational materials and experiences, while communicating a serious, impactful message, must have a certain entertaining, fun, and human quality. (This might be obvious, but upon reflection of the current state of educational materials[6] published by the city it has yet to be considered).

Through these observations we are presented with non-obvious new key questions such as what is the ideal ratio amongst levels of practitioners for the city to demonstrate a meaningful level of organizational change? How many novices to every expert is appropriate, etc.? And perhaps more interestingly: Will people adopt this without being mandated to do so?

These types of questions and synthesis in turn surface new hypothesis-driven ideas. For example, at IBM, the main incentive to "badging up" besides the implied better work output was largely social capital. As the goal is in both cases to grow skills outside of a centralized function of experts, beyond the recommendation of the CEO or trusted public figures, any mandates would have to come from smaller communities within the whole, and be a matter of personal pride.

## APPLYING SPECULATIVE FINDINGS TO A SET OF PROTOTYPES

We followed this process to re-apply insights, ideate, observe, and draw out new insights over a wide spread of original IBM research, strategies and tactics. One of the persistent challenges of speculative work is how to turn this kind of thinking into real tangible actions; for our work at A Functional Democracy, the trick is weaving this speculative thinking into a traditional human-centered-design process. For all of the scenarios and subsequent evaluations proposed through this analog, the natural next step in a design process would be to prototype small versions of each idea or repeat the cycle as needed until a compelling enough set of insights is achieved.

In the case of the badges framework, we used our Holodeck insights as a hypothesis for the future, and went about creating a prototype to test out a version of this speculative idea today.

### Civic Path and Shareable Assets

What we created first was a five level "Local Civics Path." With this framework those at a "level 0," desperate to become "level 1 Newbies" can immediately start to understand what they can do *today* to grow their practice as newly civically engaged, and also imagine a future for themselves at a more advanced level. We housed the framework within a 'zine targeted at that exact first step.
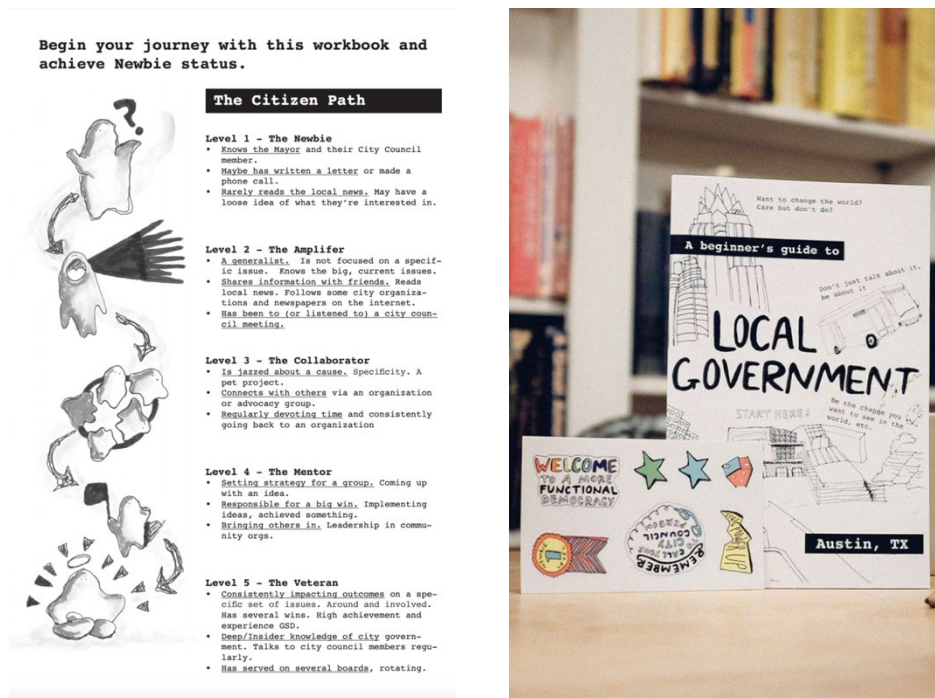


Figure 13: On the left, an excerpt from the book showing the "Civics Path" from Newbie to Mentor. On the right an image of the 'zine, "A Beginner's Guide to Local Government".

In order to create this framework, we started by sketching out draft levels based on expert interviews, and used them to source actual humans to flesh out critical behaviors and goals, further defining each level. Through this research we quickly understood that a baseline set of information (who the mayor and city manager are, how Austin's city government is structured, who your city council person is) could quickly propel people into a personal civics journey, and that that baseline was nearly invisible to most. The 'zine culminates in the "certified" levelling up for people to share socially as a marker of progress in order to empower those who self-educate and establish themselves as *Level 1 Newbies*.

We also drew from our insights at IBM in creating the framework and 'zine: we know new practitioners especially need to see the path ahead when growing a skill, with critical levels between where they are (essentially point 0) and what they see modeled by visible experts (usually levels 3+ on a 5 level scale). Importantly, they need to see this path marked by real humans with the details of personal stories to make it tangible and believable. So we included portraits of individuals at each step along the way to illustrate what it can look like to progress.



Figure 14: On the left, an excerpt from the book showing the "Level Two Amplifier" profile: Meredith. On the right an excerpt from the book showing the "Level Four Organizer" profile: Pete Rivera.

We made the book easy to produce and included stickers because we expected this scalable tool to generate excitement, and be widely shared. We made "call your Mayor" shirts

but we only sell them with the book because there is a tendency to fetishize participation with an issue such as civic engagement. Both of these insights are pulled directly from the work at IBM. The  speculative "observations" based on insights from IBM, gave us the fodder we needed to make a prototype and get started. We were able to do so with a confidence we would not otherwise have had, in a context where we had fundamentally less time, resources, and influence.

## Early Impact and Results

Our prototype is in its early phases, but in two years we've gone from a first batch of 50 'zines to over 1,500 copies distributed and sold, with 3,000 more already funded. We've used the book to make in-roads with City Hall, received multiple grants, raised thousands of dollars, and partnered with huge organizations like Google, Facebook, AIGA, Austin Park Foundation and many others, to bring "Level 1 Newbie" education to new populations. We partner with other civics engagement organizations and initiatives such as the League of Women Voters and the Workers Defense Project as channels to distribute the 'zine and its foundational education.

Here is a sample of user feedback on *A Beginner's Guide to Local Government*:

- "[I'm buying this book because I'm a small biz owner and a mom and want to make a difference." – Austin Resident
- "I am buying this book because I feel pretty lost these days with regards to government and my role in it, and feel overwhelmed about the prospect of getting involved, but also no longer feel comfortable pretending I don't need to be involved." – Austin Resident
- "I bought his book because I don't want to be passive and uneducated about this city that I deeply love and call home." – Austin Resident
- "Ladies. Thank you so much. Your Beginners Guide to Local Government is seriously perfect. You don't know how much you just inspired me and gave me hope and direction." – Austin Resident

The material has spurred a relationship with City Council, and as the leaders of the city, we know from the analog that their buy-in and promotion of this education is key to its success in terms of driving a shared narrative around the value of civic engagement and how to engage. At IBM, we see business units with strong design thinking cultures stem from both the ground up as well as executive levels—individuals who are often closely related to the DPO and help contextualize its message into their own business case. In Austin, several members of council have promoted our content, and even spoken and participated at events that surround the book and serve as next level enablement. In 2019, for example, A Functional Democracy hosted an event at City Hall for Newbies to testify in front of City Council in a special session for the first time. Over 50 people participated directly with 5 members of council including the mayor.

Figure 15: Level 1 Newbies testify at City Hall for City Council for the first time

Other prototype materials are being tested including a series of next level (Amplifier) content to help the nascent population of civically engaged continue to grow, including in-person lectures, scalable video content, and stand-alone how-to guides. A round of Civic Catalyst prototype-roles were also tested in late 2019 where 10 individuals voluntarily worked with A Functional Democracy to promote the event at City Hall and drive awareness and skills growth in their personal networks.

The work in Austin by A Functional Democracy has drawn interest from local governments and civics groups in other cities. We now have a working model for enablement and growth that can be applied (with some local expert help) elsewhere. We have learned about how to scale a localized set of practice activities to other nodes in a system from our time at IBM that can inform a fresh set of "what-ifs." The success of the prototypes to-date inspires us to continue to push the scale of the analogy, method, IBM insights, and potential growth of civic engagement outside of Austin.

## LESSONS LEARNED: EMBRACING THE VALUE OF DIFFERENCE

When we started conceptualizing IBM outcomes as ideas that could be applied to Austin, we were skeptical about the value of the comparison. What could really be similar about two groups of people that seemed so distinct? The differences between these two

organizations led to our deepest insights about human behavior and our most interesting ideas for change. Below are lessons learned from this case study.

**Organizational differences exposed insights about common human tendencies in learning and adoption of new skills.** For example, one might eagerly propose the idea of mandating the US population upskill their civics knowledge, but through reflection on IBM, we hypothesized that even if we could apply such a mandate it was unlikely to work. As a result, we were forced to jettison this easy fix idea and push ourselves to think of ways to drive adoption that involved incentivizing learning through the application to everyday life and existing social structures.

**Ideas for the future that seemed "absurd" forced us to question axiomatic beliefs about Austin's population.** As a result new potential solutions were illuminated. For example, at a corporation, a natural solution to a culture problem is to re-skill the workforce or to hire people with a new set of skills. We wondered: Could we re-skill Austin's population? Could we recruit people with civic skills to move to Austin? These were odd questions to ask about a city. What was it about a city population that made these proposals seem odd? Is it  because it's unnatural to "engineer" a population of citizens? Perhaps, but cities such as Austin frequently try to attract certain populations to drive their economies. Is it because wide-scale adult education is not perceived as a suitable role for the government? Perhaps, but cities frequently try to disseminate information to inform the residency. We found through this line of questioning that the original question was not actually the most interesting; instead, it is the line of thought that results in the question, "Why is this question strange?"

**The freedom we allowed ourselves in this process expanded the role of our ethnography from primarily tool for understanding, to tool for exploring possibilities.** When we started this project, we naturally thought to study each population independently. But through our comparison and the exciting ideas that resulted from our "what-if" scenarios, we were forced to re-imagine our ethnographic approach. What if the goal with this exercise of translating ethnographic insights wasn't to be right, but instead to investigate possibilities? This mindset shift helped us transition from exploring the world the way it is, to projecting how the world could change.

## CONCLUSION: SCALE AND THE ETHNOGRAPHY COMMUNITY

In this story ethnographic insights drive exploration into big, bold ideas about how to change our collective civic culture, starting in Austin but with potential impact for many other cities. Here the scale of impact is both about the scale of the problem, our national cultures around civic engagement, as well as the scale of the changes proposed—speculative ideas so different they seem slightly absurd. In this story, the tools of the ethnographer are scaled to a new population, to a bigger challenge and to bigger potential change.

In a world where problems are always wickeder and wickeder, analogous ethnography as a device for speculative design is just the toolset we need: it works to reveal the limitations of our imagination and the hidden structures that drive human behavior. Put in the hands of ethnographers, this approach is a way for us to explore the viability of alternative realities, gain insights into the nature of social structures, and to use those learnings to inspire systematic change in humanity's largest populations.

**Hal Wuertz** is a co-founder at A Functional Democracy and leads design teams for IBM's Artificial Intelligence Applications. Her previous paper, "Design Thinking for Change Management," outlines a sample set of ethnographic techniques used in IBM's cultural transformation. She holds a Masters of Architecture degree from Harvard University. hlwuertz@gmail.com

**Jordan Shade** is a co-founder of A Functional Democracy and leads design and human-centered strategy within IBM's services organization. She has a master of design degree from the University of the Arts, sprouting from an interest in applied design thinking and research from her work with the Reggio-Emilia pedagogy. jordan.e.shade@gmail.com

## NOTES

This paper does not represent the official position of International Business Machines Corporation.

1. For additional information on Enterprise Design Thinking, visit: ibm.com/design/thinking

2. The Forrester Total Economic Impact™ Study (2018) can be found online at: https://www.ibm.com/design/thinking/static/Enterprise-Design-Thinking-Report-8ab1e9e1622899654844a5fe1d760ed5.pdf

3. In the activity, "Big Idea Vignettes," participants use analogies to come up with unexpected ideas. (www.ibm.com/design) A similar technique is described by Ideo as "Analogous Inspiration" (https://www.designkit.org). In both cases, unexpected parallels are then used as inspiration for tangible, near-term ideas. As Ideo describes, for example, you may visit an Apple store when designing for people in difficult circumstances in order to get inspiration about how to create memorable customer experiences.

4. "Anticipatory anthropology can be variously seen as a mode of inquiry that occupies the space between the disciplines of applied anthropology and future studies. Philosophically, the anticipatory approach has deep roots in applied anthropology since the purpose of studying human experience is to improve the quality of human life in the future.... Academic or practicing anthropologists who actively consider future actions and consequences anticipate alternatives for various possible futures. These anthropologists map the implications of that flow logically or emotionally from observable practices." (English-Lueck and Avery, 2020)

5. For examples of design futures projects that leverage ethnographic techniques, see Candy and Kornet's 2019 paper, "Turning Foresight Inside Out: An Introduction to Ethnographic Experiential Futures."

6. See for example, this material published on the city's website explaining Boards and Commissions. (http://www.austintexas.gov/department/more-about-boards-and-commissions)

**REFERENCES CITED**

"Analogous Inspiration." Ideo Designkit website. Accessed [September 26 2020] https://www.designkit.org/methods/analogous-inspiration

"Big Idea Vignettes." IBM Design website. Accessed [September 26 2020] https://www.ibm.com/design/thinking/page/toolkit/activity/big-idea-vignettes

Candy, Stewart and Kornet, Kelly. 2019. "Turning Foresight Inside Out: An Introduction to Ethnographic Experiential Futures." Journal of Futures Studies website. Accessed [September 26, 2020]. https://jfsdigital.org/wp-content/uploads/2019/04/02-Candy-Turning-Foresight-Inside-Out.pdf

Currie, Adrien. 2016. "Ethnographic analogy, the comparative method, and archaeological special pleading" Science Direct website. Accessed [September 26 2020] https://www.sciencedirect.com/science/article/abs/pii/S003936811500103X

Dunne, Anthony and Raby, Fiona. 2013. Speculative Everything: Design, Dreaming and Social Dreaming. MIT Press.

English-Lueck, J.A. and Avery, Miriam. 2020. "Futures Research in Anticipatory Anthropology." Oxford Research Encyclopedias website. Accessed [September 26, 2020] https://oxfordre.com/anthropology/view/10.1093/acrefore/9780190854584.001.0001/acrefore-9780190854584-e-14

"Enterprise Design Thinking." IBM Design website. Accessed [September 26, 2020] ibm.com/design/thinking

The Forrester Total Economic Impact™ Study (2018) can be found online at: https://www.ibm.com/design/thinking/static/Enterprise-Design-Thinking-Report-8ab1e9e1622899654844a5fe1d760ed5.pdf

CATALYST

# Scaling Out (Not Only Up)

## Distributed Collaboration Models to Get Work Done

ALICIA DORNADIC, *MindSpark*
NIKKI LAVOIE, *MindSpark*
SHEILA SUAREZ DE FLORES, *10 10 10, Eco.tter*
ELVIN TUYGAN, *MindSpark*

*In this catalyst, we the authors describe the benefits of 'scaling out': reaching out beyond one's organization to bring in external partners to accomplish UX research. Organizations scale out their research efforts in order to cover more ground, draw from more specialties, or conduct more research more quickly than they would be able to alone. As opposed to growing an in-house team to meeting research needs ('scaling up'), scaling out can be a more inclusive approach to generating user insights, where the voices of diverse research partners throughout the world are brought together to produce powerful UX research outcomes. A case study example of work with suppliers and clients illustrates scaling out. Collective intelligence pushes scaling out even further, as it counts research participants, users and potential users as part of the network of partners who get work done.*

*Keywords: distributed collaboration, scaling out, collective intelligence, global research*

## INTRODUCTION

For many years, we at EPIC had to argue for the value of ethnography and of user research, more generally inside our organizations. Now that we have won a seat at the table, we are able to create stronger networks internally within our own organizations and externally, with partners outside. Recently there has been a push for 'scaling up' research teams within organizations. We have learned from this that it is no easy feat to go from a research team of two to thirteen (Clancy 2018); or to adapt the role of UX research as the team scales (Primadani 2019); or to prove inside organizations the value of research, while implementing processes of working cross-functionally, and to manage the actual scaling process simultaneously (Chokshi 2019).

In this paper we make the argument for the benefits of scaling out. We define 'scaling out' as reaching out beyond one's organization to bring in external partners to accomplish UX research, cover more ground, to bring in more specialties and ultimately rely on collective intelligence to get work done. Scaling out is an inclusive approach that can produce meaningful results due to presence of diverse voices, including among the UX researchers working on the project. Today with more social networking and digital collaboration tools than ever before, as we describe in this Catalyst, scaling out has become manageable and efficient. We also consider power imbalances in UX research, particularly in the roles of client-supplier, researcher-participant and designer-user, and whether it is possible to scale out to such a point that we shift them. Can we as UX researchers, go from evangelizing user research and amplifying the voice of the user to collaborating continually

with users - with people. We call upon the EPIC community to consider what it means to be on one side of the researcher-participant dynamic and ways to shift this while working in the realities of market economies and for-profit business sectors.

We believe this conversation on scaling out is more timely than ever in a COVID-19 world. The first half of this Catalyst tackles the practical side of today's restrictions. As the conditions which make possible in-person research are tenuous or in flux (physical space restrictions such as lockdown, curfew, masks), we must rely more than ever on distributed local researchers, participants, and digital collaboration tools. The EPIC community will already be familiar with some of these tools and their advantages.

In the second half of the Catalyst, we take the concept of 'scaling out' to its logical conclusion. We speculate on a form of distributed UX research where our global partners are not local professionals, but rather "citizen researchers," or even users themselves. We imagine scenarios where user research is conducted by crowdsourcing and collective intelligence, similar to models like Wikipedia and some COVID-19 response projects (OECD.org). This is not a new phenomenon but we believe it is more timely than ever.

## BACKGROUND

EPIC practitioners have written about the obstacles and opportunities of globally distributed teams and partnership through ethnography. From the early days of EPIC, Mack and Mehta (2005) describe, "as more and more corporate ethnographic work is crossing international borders, we are increasingly collaborating with teams that are spread across the globe." Churchill and Whalen (2005) emphasize the methodological challenges of geographically distributed projects. These include the extra effort required to build and maintain relations and empathy; widely varying experiences of ethnographic methods, local language and culture; and conflicting responsibilities and lines of accountability. At times, researchers train non-researchers to help conduct research. Di Leone and Edwards (2010) point to four key needs for knowledge sharing in collaborative ethnographic research. Brannen, Moore, and Mughan (2013) describe a project in which they acted as outside experts in collaboration with a multicultural project team of nine managers within an organization. The managers were trained in ethnographic techniques, and they conducted the data collection and participated fully in analysis. The consultants acted as advisors, trainers and coaches to the project team at every stage of the ethnographic process. Kearon and Earls (2010) describe a project that employed participants to conduct ethnography themselves where the initial batch of results were lackluster. They urge careful framing and training in order to obtain useful results.

As remote research and collaboration tools advance, so do the conversations at EPIC. Gorkovenko et al.'s paper (2019) describes a project in which ethnographers and participants were able to engage remotely in contextual inquiry around a product with the aid of sensing technology. Golias (2017) discusses how ethnography and remote usability testing can enhance one another.

Today, as travel restrictions loom due to COVID-19, these conversations take on a new sense of urgency. Henshall (2020) describes the benefits of remote research including participants being comfortable in their home environment, feeling safer sharing their point of view in a private space, reaching a larger pool of participants, scheduling flexibility, and including more stakeholders via remote observation. Drawbacks being that it is difficult to

include unconnected households, people less comfortable with digital tools who then require extra support to participate, finding a private spot in a crowded household, and framing being limited. The researcher's ability to zoom into details or zoom out for a broader view is not always possible. Evans (2020) echoes many of these sentiments when describing her own remote ethnographic research, relying heavily on diary study tools, online surveys, and participant-made videos. She offers solutions to overcome obstacles of remote research, including finding the least data heavy video conferencing tools, becoming more prescriptive in questions for participant-made videos, and balancing these data with other sources of insights. In a recapitulation of an EPIC panel from May 2020, Collier Jennings and Denny (2020) paint a picture of how ethnography evolves and how research today weaves together many different vantage points. If one view is eliminated or limited (such as, conducting in-person research) how can we use other available data points to provide a holistic picture? They also describe participant- and community-led engagement and co-creating insights.

## SCALING OUT: A CASE STUDY IN DISTRIBUTED RESEARCH

Our Paris-based boutique agency, MindSpark, relies on scaling out our resources to accomplish projects. We offer UX and market research on a global scale to meet our clients' research objectives. By scaling out we mean that we work with suppliers and clients to design and execute research by bringing in stakeholders and partners with specialized skillsets, expert and/or localized knowledge, as they are needed. Scaling out allows an organization broad reach and flexibility by relying on local experts globally.

For example we, the authors, worked on a project for a large tech company for which we conducted a user experience research study in five markets: South Korea, Bangladesh, Chile, Kenya, and Thailand. The company approached us to design the research, to collect data, and to deliver findings with our local partners. Before the client came to us, it would have been difficult to anticipate a need for resources in these markets, both internally for the client or for us as a research agency. The client did not have researchers on the ground in those countries, nor the network to contract directly with local researchers. We did not have in-house researchers on the ground full-time in those markets.

What we *did* have to offer was an extensive, trusted network of suppliers with whom we collaborate and could quickly enlist. When conducting user research on a global product, it often becomes necessary to "pick up" partners when a research question arises in a corner of the world little known to the researcher. We first piloted the research in Korea and the core team including two MindSpark researchers and four clients traveled to Seoul to observe and participate in-person alongside the local team, consisting of one moderator, one project manager, and a simultaneous translator. Then we, from MindSpark with the clients, worked remotely with local teams to execute the other four markets concurrently. By letting go of total ownership of this project and relying on the expertise of local researchers, we were able to gain insight into how people in those markets would want to interact with our clients' product. We were also able to save time by running the studies concurrently.

When we scale out at MindSpark, we intentionally keep the research design flexible enough to allow local partners to tailor the methodology to their localized context. In the same study, the local Kenya team recommended conducting interviews in-person (pre-COVID-19) at local offices or cafés. In Chile, internet connection is strong enough it was deemed most appropriate to conduct remote interviews to capture a more geographically

dispersed audience. The local researchers were able to provide much more cultural context than as outsiders we would have been able to glean on our own. Themes around public/private spaces, security/protection, sexuality and violence (all related to the product) emerged in conversations with the local researchers.

We often collaborate with researchers with special skillsets. In this project, we were able to work with moderators who has specific experience in the topic matter of the study and were able to provide additional insight and background knowledge for the final reports. In other instances, we might rely on a quant guru to validate or supplement the qualitative pieces we would be working on. We collaborate with colleagues specialized in visual design, UI and UX design, creative agencies who might be delivering assets that we could then test through research, inspiring creative production or testing it afterwards.

## TOOLS OF THE TRADE

There are a few basic elements to MindSpark's model of scaling out. First, we establish a core team, usually (at the very least) one researcher, one project manager, and one client point person who will see a project through for a certain period of time, usually from start to finish. This gives consistency and rigor to the work. Second, as mentioned above, we gather a network of partners with specialized skillsets or local knowledge. The core team acts as a hub that gathers and transmits information to that network. Third, we create extra layers of communication and documentation. This can include phone or video calls to review objectives and to introduce all the main players of the project to each other. This can mean separate project messaging channels (on Slack or WhatsApp, for example). It can take the form of shared calendars, google documents, collaboration boards (on Notion, Mural, Miro etc.) It means documenting objectives from beginning stages to end deliverables. When working remotely or in distributed teams, extra emphasis needs to be placed on careful communication in order to maintain shared goals and contribute meaningfully to the same output. Summarizing advances or changes in the project in emails or shared digital document spaces (Dropbox, Google docs) is crucial. To this end, templates are extremely helpful. Templates help ensure logic and consistency to note-taking, findings, and final reporting.

Ongoing conversations can occur around language choice and cultural context, particularly when testing copy or content. Cross-cultural conversations can be around screener crafting and recruitment. Perhaps the target participant is not representative or might not even exist in a particular market. Legal limitations differ across regions that will impact a research study. Even budgeting time per interview can be an issue. A user interview that might take 60-minutes in the US could run closer to 75 minutes in a different market. We usually require at least 30-45 minute buffers between remote interviews in markets where bandwidth is unreliable or where people tend to share computers and other devices and they might not be familiar with troubleshooting those particular devices.

## BENEFITS OF SCALING OUT

A first benefit of scaling out is the ability to incorporate deep local knowledge into the research in order to produce a hybrid insider-outside view. By partnering with local teams, the researcher gets to come in as an outsider with a naive understanding of local conditions.

The local moderator has the language and cultural background to execute the study and to assist greatly with analysis and findings generation.

By partnering, the researcher can become the hub, with a higher systems-level view. This gives us the ability to cut across organizational silos, across different markets, and across different populations of users. In another project, we worked with stakeholders from three different departments across two organizations, connected by a particular user journey we were charged with mapping. By using an ethnographic approach to study the whole system and to present that system back to stakeholders, we were able to help the organizations make strategic decisions to improve their internal processes and, ultimately, the user experience. Our advice could be considered as having less bias, due to not being part of the structure and their system of bonuses and promotions.

Another benefit is that it is less risky for us to try new approaches as outsiders. We can break the mold and experiment in ways that might be more difficult to do in-house or without a partner network. We have the advantage of multiple view points grounded in local contexts. They will be more relevant, diverse, and simply having more brains on a project with new ideas can help challenge our preconceived ideas of how a project should run.

Scaling out also has the benefit of budget precision, executing research and hiring researchers when and where they are needed, bringing in necessary skillsets or expertise in a certain geography, language, subject to answer particular questions. It has the advantage of speed, particularly when running multiple market studies concurrently. Arguably, the outcomes will be of higher quality when produced by a series of collaborators rather than a sole researcher attempting to bring all these pieces of data together alone. The quality will be enriched due to the inherent diverse nature of on the ground voices, lived experiences, and styles of approach. In the next session, we speculate on how deep this democratization of research could go in the future.

## CONSIDERING COLLECTIVE INTELLIGENCE

### Story Time: Multiple Viewpoints

Imagine a particularly enterprising traveling ant trying to inspect an elephant, an animal it has never seen before. But the elephant is too big to see from such a small viewpoint. The ant may only see a hoof and think the elephant is in fact a rock, something it has seen before. But what if this ant brought more ants to crawl around, under and on top of the elephant to report back and create a 360-degree view? If it multiplies its viewpoints it can not only minimize movement but save time and spend more time observing from its chosen spot. "Oh, wow, this rock moves from time to time! Danger!" Better yet, this upstart ant could ask a local anthill for help. They know their elephants!

Figure 1. Elephant and Ant, cartoon image, 1948, [commons.wikimedia.org](commons.wikimedia.org) public domain

At MindSpark, we are continually pushing our projects to be more inclusive, more collaborative, and to bring clients, partners, and participants together in meaningful ways. We look to collective intelligence as one potential model for expanding the roles, responsibilities and voices that participants bring to projects. What could it look like if we narrowed the distinction between researcher and subject, and subjects became simply informant-researchers reporting from everywhere, particularly in this moment in time when digital tools are widely available and travel restrictions and social distancing orders prevent researchers from going to multiple markets to conduct in-person research? Collective intelligence is a distributed versus centralized model for advancing towards a goal (Lévy 1981). It relies on diversity of view point, motivation, lived-experience and knowledge, auto-organisation and convergence of ideas of actors in an existing or newly formed network, community or subset of the population.

The quality of the collective intelligence can be enhanced by ensuring as much of a flat hierarchical structure as possible (including relations with partners and reduced contract leverage pressure), lean process controls that do not throttle input or research style, or encouragement of self-organization by different parties. Collective intelligence is not homogenization and delegation of directives (how and what) to multiple parties but rather empowerment, collaboration and trust in individual skills and abilities to create their own how and what before it merges back to aid the common goal. If the end project design is then implemented by those in the collective intelligence project, they are more likely to fully implement and grow the following product (Nguyen et al 2019).

As applied to UX research, perhaps a simple example of this is conducting research or collecting feedback about a product in the precise moment when user motivation is high. In

the moment when, let's say, a person is looking at real estate listings online and is struggling with the user interface, for example. They are motivated to improve the interface or to even provide context around their search. Bolt and Tulathimutte (2010) describe the importance of recruiting for moments of motivation.

We can flip the intention of many usability and user tests. Rather than testing whether a designed product is desirable and usable by a target population, we can design products to be imbued with meaning, improved and challenged by the very people who have the interest in using it. We can be utilizing open source tools for data collection and to produce insights. For example, Stamen Design, a data visualization and cartography studio in San Francisco offers tools and visualizations that are meant to be picked up and improved by the people they are representing. One of their projects, Field Papers which launched in 2012, "allows people to create a multi-page atlas of anywhere in the world. Once you print it, you can take it outside into the field, record notes and observations about the area you're looking at, or use it as your own personal tour guide in a new city."

## Collective Intelligence in Action: Crowdsourcing
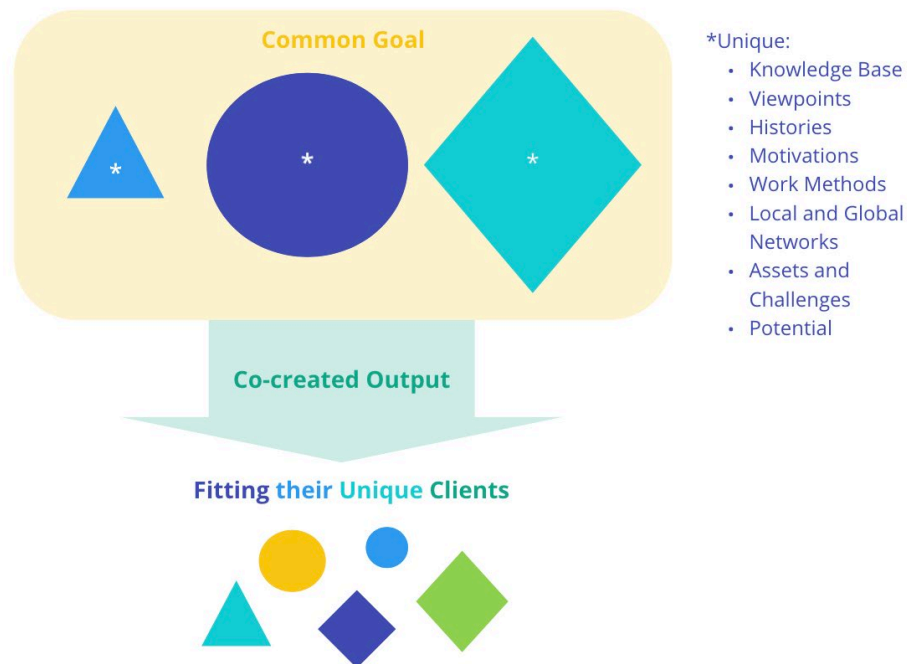


Figure 2. Collective intelligence as applied to UX projects in which clients, suppliers and participants work together to produce knowledge that eventually influences product-service offerings. © Sheila Suarez de Flores.

If we consider our participants to be collaborators, can we imagine if we not only scaled out our internal teams to include external partners, and to include users, but also to include the whole world of citizens happily collaborating on our projects?

Crowdsourcing is "the act of taking a job traditionally performed by a designated agent and outsourcing it to an undefined, generally large group of people in the form of an open call" (Howe 2008). Currently this is already happening in open initiatives such as openCovid19 and others across the globe to tackle COVID-19 related challenges. We would be remiss to not mention Wikipedia which is arguably the most well-known example of collective intelligence (Malone, Laubacher, Dellarocas 2009). NASA is collaborating with "citizen scientists" in various projects such as landslide reporting (Cooperative Open Online Landslide Repository), comet discovery (The Sungrazer Project) and even finding a new planet or "Planet 9" (a project with over 62,000 participants).

This may all seem a bit futuristic but in fact crowdsourcing or other open collaborations across "organizations" is quite ancient. Even trees use it! In Peter Wohleben's book (2016), *The Hidden Life of Trees: What They Feel, How They Communicate*, he explains that trees communicate through the air, using pheromones and other scent signals. Such as in sub-Saharan Africa, the wide-crowned umbrella thorn acacia will emit a distress signal scent when being chewed upon to warn neighboring trees. These trees then change their leaf composition to become deadly to herbivores. Luckily (or unluckily for trees) herbivores, such as giraffes, have learned to eat down wind (Grant 2018).

## Creating Citizen Researchers

In order to start harnessing collective intelligence, research organizations might typically examine existing organic data artifacts from the customer community, or widen their data stream to include inputs from external partners. But collective intelligence is not fully realized without a group of human actors working together toward some common goal and/or framework. To that end, we propose creating "citizen researchers": a pool of citizens (including potentially customers) engaged and motivated to work towards finding a research result or solution for a target user.

Citizen researchers would do for UX research what citizen scientists have done for environmental science since at least the 1990s, when the term was coined, if not for millennia. Alan Irwin, a sociologist now based at the Copenhagen Business School, defined citizen science both as "science which assists the needs and concerns of citizens" and as "a form of science developed and enacted by the citizens themselves" (Irwin 2018).

A paper by Nguyen (2019) on the practice of developing citizen research projects outlines these key steps to begin:

1. Identify the research question and the communities of participants.
2. Decide on incentives to engage participants: "a combination of both extrinsic motivators such as authorship and access to the data and intrinsic motivators such as making tasks enjoyable, offering participants the opportunity to gain new knowledge and finding meaningful outlets for their skills."
3. Determining methods to evaluate solutions created by collective intelligence and decision making.

It is important to know that to get optimum results from the above process, and despite conventional wisdom, the reflex should be to include everyone (citizen researchers and all partners) as much as possible through the entire project process including setup and synthesis. The efficiency gained is in the bespoke output unique to the diverse parties that create it and not necessarily in the reduction of interactions.

Be aware of the balance between evaluation and pure creation when we let diverse voices collaborate in order not to filter out pertinent results, or representative minorities. Once the above project foundation is complete, the team (internal/external) still should adjust it in iterations based on the feedback of included citizen researchers. The next steps are to select and onboard a small subset of representative citizen researchers that will help complete the project structuration, solidify the goal, and set up the framework or platform for input and tasks by crowds collaborating with you. A clear ongoing communication plan is also key to onboard and keep everyone engaged (Nguyen et al 2019).

## Use Cases for Citizen Researchers

Like all solutions there is never, and should never, be a one-size-fits-all framework. Here are some use-cases where 'scaling out' is expanded to include citizen researchers.

1. To study hard-to-reach populations. For example, you want to approach an elderly population with unstable access to the internet, and who are hard to travel to. One solution is to form a team of citizen researchers (probably composed of concerned citizens or more connected relatives) to help, with permission of the target subjects, document parts of the daily lives of the target subjects.

2. To create loyalty and foster innovation. For example, a business could create a platform for citizen researchers to share parts of their lives (permission required) using a product or service, with the incentive of joining a community and improving a product and/or service they care about. Within this platform, there could be a system of suggestion and upvote much like on Reddit or Threadless to collaborate on the synthesis and solution emergence mode.

3. To ensure buy-in of proposed solutions. This may be the case especially for social innovation research where lives are trying to be saved or improved, such as building public spaces and services (e.g. health campaigns). Not only should the citizens that will be using the solution be invited to collaborate but the recognized leaders (official or not) (Nygun et al 2019). This is also called the Ikea Effect, where customers often are more attached and likely to use an item if they made it themselves (Norton, Mochon, Ariely 2011). Participatory design, an approach that includes a range stakeholders (employees, customers, users) in the design process to ensure the end results meet their needs, is not new. However, again due to social media and digital tools, the co-production process can be diversified and enhanced today (Devisch, Huybrechts, De Ridder 2019).

## COLLECTIVE INTELLIGENCE: AVOIDING ETHICAL TRAPS

Once you start seeing collective intelligence it is hard to stop seeing it all around you. It is everywhere in many different scales, cross-sections and flavors. The cells in our body and

bacteria in our gut are even considered a form of collective intelligence (see: big brain or swarm intelligence). Bees do it! Birds do it! Computers do it (AI)! And now UX researchers can do it even more, especially due to technology. We cannot help but hypothesize that this also has the hidden benefit of promoting research best practices and critical thinking. A whole new world of researchers!

Still there are some ethical traps to avoid. The key one is the use and abuse of our citizen researchers if they treated as "cheaper labor". We do not want to turn them over to the gig-economy where collaborators, often from marginalized or lower income communities are reliant on incentives and "fired" at whim. Gray and Suri (2019) describe the invisible labor that powers Silicon Valley, such as manual image recognition and data sorting.

We as researchers can avoid creating an unethical power balance by properly paying and ensuring longevity of contract and support at the end when citizen researchers create this as their full or even partial time gig. Otherwise, monetary rewards could be avoided while leaning into other extrinsic or intrinsic rewards such as esteem and community (e.g. NASA).

Another ethical trap is any sort of research or collective intelligence analysis done on data where the respondents are not fully aware of the treatment of their data or its goal. It may be tempting to create a platform where people share about their day or other people's with a simple use and terms statement that participants click through without reading. It is advised to not only make the project clear from sign-up, including the goal and purpose in main marketing and how the data will be used.

The last and most important tip is for "researchers using collective intelligence to show their results…evaluate… and be transparent about mistakes." Rigorous evaluation of collective intelligence is necessary to provide evidence of its usefulness to stakeholders, "so that it gets recognised and funded properly" (Nguyen et al 2019). But we would add, also to be able to adjust and learn through the process, create loyalty and reaction from participants. And, most importantly, at the core why we do it, in order to best serve our citizen researchers and target users.

## WHERE DO WE GO FROM HERE?

Scaling out can be done to degrees. Scaling out can mean bringing in external consultants to execute a project. It can include distributed teams within an organization. It can be webs of networked partners that assist at various points in a project for different purposes (various geographic regions, language, expertise). It can be pushed further to include users, customers, and everyday "citizens" as creators and sources of data.

There are certainly challenges to scaling out. In our MindSpark model, we have found that challenges can arise around ownership of proprietary or internal processes. There can be procurement challenges of all these external resources and it can be less predictable or manageable how all these external sources will come to participate in one organization's or one team's research plan. More effort to communicate, with the aid of digital tools, is needed when working externally and across distributed groups. For those with concerns of data ownership and confidentiality agreements, this legal aspect is a bit more complicated. Not all data can be shared openly. There are privacy and confidentiality restrictions to be considered.

The future is unknown, it would be foolish to say with any certainty how collaboration will take shape in these times of COIVID and/when/if Post-COVID-19. But what seems

certain is we should make an effort to try new forms of partnership globally and to take advantage of ones that already exist but in other contexts. As many of us are working remotely, with travel restrictions, it seems logical to tap into distributed knowledge sources. This can be an opportunity to rethink our relationships to internal and external partners, and to our participants. It can be an opportunity to rethink the who even can be a participant or researcher. We can imagine distributed collaboration models that include the user/participant, and shift the power imbalance of researcher-participant by giving participants a seat at the table and engendering authentic empathy through continued interaction.

**Alicia Dornadic** is a design anthropologist at MindSpark. She has ten years of experience in user research in the realms of work, health, social media, transportation, and public safety. alicia@mindsparklab.com

**Nikki Lavoie**, Founder of MindSpark Research International, is a spirited and intuitive qualitative UX researcher who translates her passion for understanding people into strategic insight. She has focused on combining ethnographic and digital techniques during her years spent in the US and now as an expat living in Paris.

**Sheila Suarez de Flores** is a product design and business transformation coach, passionate about ecological and social responsibility, circular design and systems thinking. She has experience in various sectors and sized businesses (local to global). She is an expert in agile, lean, distributed teams, workshop creation, facilitation, and collective intelligence. me@sdeflores.com

**Elvin Tuygan** is a senior design anthropologist at MindSpark with over 15 years of global strategic experience, passionate about all things digital, from gaming to social networking. She has studied, lived and worked around the world, including Boston, Rabat, London and Istanbul and has now settled in the South of France.  elvin@mindsparklab.com

## REFERENCES CITED

Bolt, Nate and Tony Tulathimutte. 2010. *Remote Research: Real Users, Real Time, Real Research*. Brooklyn: Rosenfeld Media.

Brannen, Mary Y., Fiona Moore, and Terry Mughan. 2013. "Strategic Ethnography and Reinvigorating Tesco Plc: Leveraging Inside/Out Bicultural Bridging in Multicultural Teams." *Ethnographic Praxis in Industry Conference Proceedings:* 282-299.

Chokshi, Monal. "How to Succeed as a UXR Manager: Capitalizing on your Strengths as a Researcher." *Strive: The 2019 UX Research Conference.*

Churchill, Elizabeth F. and Jack Whalen. 2005. "Ethnography and Process Change in Organizations: Methodological Challenges in a Cross-cultural, Bilingual, Geographically Distributed Corporate Project." *Ethnographic Praxis in Industry Conference Proceedings:* 179-187.

Clancy, Charlotte. 2018. "Scaling the Deliveroo User Research Team." Accessed September 22, 2020. https://medium.com/deliveroo-design/elevating-the-impact-of-our-research-3267068937fe

Collier Jennings, Jennifer and Rita Denny. 2020. "Where is Remote Research? Ethnographic Positioning in Shifting Spaces." *Ethnographic Praxis in Industry Conference Perspectives.* Accessed, September 22, 2020. https://www.epicpeople.org/where-is-remote-research-ethnographic-positioning-in-shifting-spaces/

Cooperative Open Online Landslide Repository. NASA. https://gpm.nasa.gov/landslides/ Devisch, Oswald, Liesbeth Huybrechts, and Roel De Ridder. 2019. *Participatory Design Theory: Using Technology and Social Media to Foster Civic Engagement.* New York: Routledge.

Di Leone, Beth and Elizabeth Edwards. 2010. "Innovation in Collaboration: Using an Internet-Based Research Tool as a New Way to Share Ethnographic Knowledge." *Ethnographic Praxis in Industry Conference Proceedings*: 122-135.

Evans, Chloe. 2020. "Ethnographic Research in Remote Spaces: Overcoming Practical Obstacles and Embracing Change." *Ethnographic Praxis in Industry Conference Perspectives.* Accessed September 23, 2020. https://www.epicpeople.org/ethnographic-research-in-remote-spaces-overcoming-practical-obstacles-and-embracing-change/

Golias, Christopher A. 2017. "The Ethnographer's Spyglass: Insights and Distortions from Remote Usability Testing." *Ethnographic Praxis in Industry Conference Proceedings:* 247-261.

Gorkovenko, Katerina, Dan Burnett, Dave Murray-Rust, James Thorp, and Daniel Richards. 2019. "Supporting Real-Time Contextual Inquiry through Sensor Data." *Ethnographic Praxis in Industry Conference Proceedings*: 554-581.

Grant, Richard. 2018. "Do Trees Talk to Each Other?" *Smithsonian Magazine.* Accessed September 22, 2020. https://www.smithsonianmag.com/science-nature/the-whispering-trees-180968084/

Gray, Mary and Siddharth Suri. 2019. *Ghost Work: How to Stop Silicon Valley from Building a New Global Underclass.* New York: Houghton Mifflin Harcourt.

Henshall, Stuart. 2020. "Recalibrating UX Labs in the Covid-19 Era." *Ethnographic Praxis in Industry Conference Perspectives.* Accessed September 23, 2020. https://www.epicpeople.org/recalibrating-ux-labs-in-the-covid-19-era/

Howe, Jeff. 2008. Crowdsourcing: *Why the Power of the Crowd is Driving the Future of Busines*s. New York: Crown Business.

Irwin, Aisling. 2018. "No PhDs Needed: How Citizen Science is Transforming Research." *Nature.* Accessed September 22, 2020. https://www.nature.com/articles/d41586-018-07106-5#:~:text=The%20phrase%20'citizen%20science'%20itself,enacted%20by%20the%20citizens%20themselves%E2%80%9D.

Just One Giant Lab. Open COVID19 Initiative. https://app.jogl.io/program/opencovid19

Kearon, John and Mark Earls. 2009. "Me-to-we Research: Digital Characters." *ESOMAR Congress 2009: Leading the Way.*

Lévy, Pierre. 1981. *L'intelligence collective: Pour une anthropologie du cyberspace.* Paris: Éditions François Maspéro.

Mack, Alexandra and Dina Mehta. 2005. "Accelerating Collaboration with Social Tools." *Ethnographic Praxis in Industry Conference Proceedings:* 146-152.

Malone, Thomas W., Robert Laubacher, and Chrysanthos Dellarocas. 2009. "Mapping the Genome of Collective Intellgience." *MIT Center for Collective Intelligence.* Accessed September 22, 2020. http://www.realtechsupport.org/UB/MRIII/papers/CollectiveIntelligence/MIT_CollectiveIntelligence2009.pdf

Nguyen, Van Thu, Bridget Young, Phillippe Ravaud, Nivantha Naidoo, Mehdi Benchoufi, and Isabelle Boutron. 2019. "Overcoming Barriers to Mobilizing Collective Intelligence in Research: Qualitative Study of Researchers With Experience of Collective Intelligence." *J Med Internet Res.* Vol 21. Iss. 7. 8-9. Accessed September 22, 2020. https://www.academia.edu/39881568/Overcoming_Barriers_to_Mobilizing_Collective_Intelligence_in_Research_Qualitative_Study_of_Researchers_With_Experience_of_Collective_Intelligence

Norton, Michael, I., Daniel Mochon, and Dan Ariely. 2011. "Ikea Effect: When Labor Leads to Love." Science Direct. Accessed September 22, 2020. https://www.sciencedirect.com/science/article/abs/pii/S1057740811000829

OECD Policy Responses to Coronavirus (COVID-19). 2020. "Crowdsourcing STI policy solutions to COVID-19." Accessed September 22, 2020. http://www.oecd.org/coronavirus/policy-responses/crowdsourcing-sti-policy-solutions-to-covid-19-c4f057b3/

Primadani, Luky. 2019. "Strategies in adopting and scaling user research within 1.5 year." *UX Collective.* Accessed September 22, 2020. https://uxdesign.cc/strategies-in-adopting-and-scaling-user-research-within-1-5-year-ad7ca28b3490

Stamen Design. *Field Papers.* Accessed September 23, 2020. https://stamen.com/work/field-papers/

Wohleben, Peter. 2016. *The Hidden Life of Trees: What They Feel, How They Communicate - Discoveries from a Secret World.* Vancouver: Greystone Books.

**THEMATIC SESSION**

# Post-Human Scale

In this session we examine the interfaces of humans, software, data, and machines. We'll consider how ethnography helps us understand beyond the human scale, how we need to adapt our practices to explore these relationships, and how ethnographic insights can be fed back into the more-than-human systems we build.

Session Curators: Laura Cesafsky, Kat Ekberg, Evan Hanover, Scott Matter

CASE STUDY

# Software Quality and Its Entanglements in Practice

JULIA PRIOR, *University of Technology Sydney*
JOHN LEANEY, *University of Technology Sydney*

*Effective software quality assurance in large-scale, complex software systems is one of the most vexed issues in software engineering, and, it is becoming ever more challenging.   Software quality and its assurance is part of software development practice, a messy, complicated and constantly shifting human endeavor.*

*What emerged from our immersive study in a large Australian software development company is that software quality in practice is inextricably entangled with the phenomena of productivity, time, infrastructure and human practice. This ethnographic insight --- made visible to the organization and its developers via the rich picture and the concept of entanglements--- built their trust in our work and expertise.  This led to us being invited to work with the software development teams on areas for change and improvement and moving to a participatory and leading role in organizational change.*

*Keywords:* ethnography, entanglements, rich_pictures, software_development

## INTRODUCTION AND CONTEXT

Effective software quality assurance in large-scale, complex software systems is one of the most vexed issues in software engineering.  Today's software systems provide sophisticated functionality that was not even imaginable a couple of decades ago – and assuring quality in these increasingly capable, adaptive and connected systems is becoming ever more challenging (Mistrik et al. 2016).  Software quality and its assurance is part of software development practice – a messy, complicated and constantly shifting human endeavor.

We were drawn to the participant organization – a software development company – as a fascinating and promising place to explore challenges in quality assurance through the lived experience of  professional software developers. Its flagship software product,  Connect (a pseudonym), is an extraordinarily large and complex software system used by thousands of customers in dozens of countries across the globe.  The organization's profound knowledge and experience in the industry it has served for over two decades, and a continual development approach, ensures that Connect's functionality becomes more and more advanced every year.

At the time of the fieldwork, there were just over two hundred software developers working collaboratively in a dozen software product development teams. They were based primarily in the head office in Australia, with most responsible for different modules of functionality in Connect, and a couple of smaller teams developing separate products that interacted directly with Connect.

Keeping Connect performing reliably for its tens of thousands of users necessitates robust software development, quality assurance and work management processes.  The company has developed a high quality ethic around its development of software over many years. This has come about by the continuing discussions around quality and productivity

that permeate all teams. The quality assurance processes it has in place make for an extremely robust product that is recognised as such by the industry that it serves (Prior 2011). Nonetheless, challenges to the reliability of these quality processes were posed by the consistent growth of both Connect and the number of new developers unfamiliar with the organization and its complex systems, processes and practices.

This case study is based on the ethnographic work we carried out when the first author spent a six-month sabbatical working full-time in the organization.

## RICH PICTURES

As this is ethnographic work, there is of course thick, rich data. We needed to make friends with all of this data, to manage and analyse it without becoming overwhelmed. It became apparent as data was collected that some diagrammatic means of representing the relationships discovered from analysis was important. For a large system, or complex environment, diagrams "encourage holistic rather than reductionist thinking about a situation" (Checkland 2000).

Rich pictures are compilations of drawings, pictures, symbols and text, that show relationships, connections, influence, processes, as well as characters and characteristics, points of view, prejudices and preferences.

We chose Checkland's rich pictures as they are not hierarchical, can be used to extend analysis via the Soft Systems Methodology (SSM) and both authors were familiar with the SSM and had used it in the past. SSM shares similar theoretical underpinnings to an ethnographic approach. The most notable one is the lack of belief in a universal theory, or driving system, for an organization.

The rich pictures we created with developers proved particularly useful for:

- Exploring and identifying aspects and perspectives to include in mapping a system or situation
- Capturing structure and process of what is happening in a situation, as well as people's feelings, values and perspectives
- Fostering communication with others about a situation
- Developing a shared understanding of a situation or initiative as a group
- Motivating further discussion, learning and/or action
- The unanticipated effect of the rich picture was the deeper engagement of the developers with our work, which helped build their trust in us, and appreciation of our research.

## BUILDING A RICH PICTURE OF THE ORGANISATION

### A Rich Picture

The first author started the rich picture by using Post-Its, as she could place and move them around easily.

This piece of the first, rather rough, rich picture represents activity around the testing process: we see a couple of new developers writing some unit tests and using the automated test system, DAT. We also see their code going through a couple of iterations to improve quality. These iterations are about the code being refined, and eventually the code being checked into the code database. In their development process, TestFirst is a fundamental
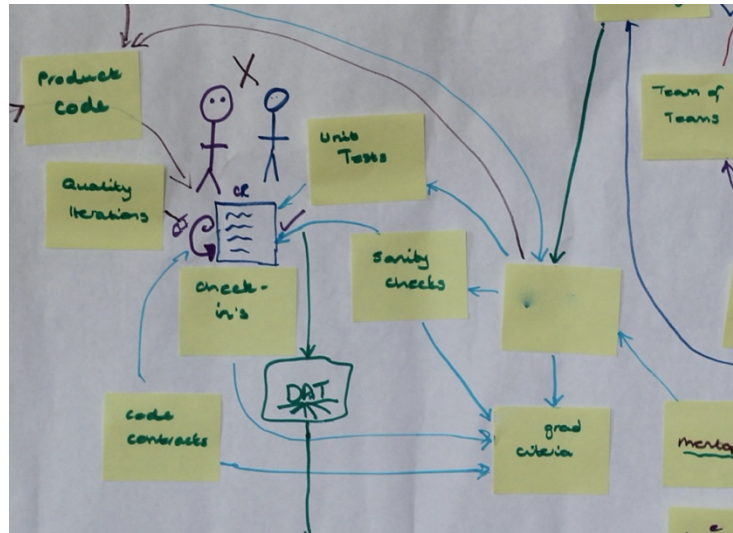


Figure 1. First rich picture composed of Post-its.

approach in which the tests are written before the functional code. It directs thinking towards outcomes, and how they will be tested to demonstrate correctness.

Our observations were that the more experienced developers will talk about the essential use of TestFirst – as a design approach, but also for investigating and fixing defects:

"Let's write the test first, and then see if we need to change the others [unit tests]."

In the picture, we see the developers performing tests, driven by quality needs. In tension, they are also driven by the need for progress, as expressed by the Post-It labelled 'check-in's'. A check-in occurs when a developer uploads their final tested and peer-reviewed new or revised code to the main codebase. One can also see the interactions with new and trainee developers in the company, shown by the Post-Its, 'sanity checks', 'grad criteria' and the redacted Post-It. A 'sanity check' is a brief run through of the functionality of the code to establish that it works more or less as expected; 'grad criteria' refers to the set of measures that a new developer must meet before they can graduate from, or complete, their probationary training – these include a minimum number of check-in's and sanity checks, for example.

For many developers, these training interactions cause tension in achieving productivity, as demonstrated by the following quote:

"All the senior devs. are already busy doing what they are working on at the moment. It's kind of like, they've got their work and they have to teach other people at the same time. So the priority for senior dev. is, of course, their current work."

In summary, what we were seeing more clearly via the Rich Picture was an understanding that was broader (more of the interactions within and between teams and the influences on developers' behavior) and deeper (more subtle interactions).

**A Richer Rich Picture**
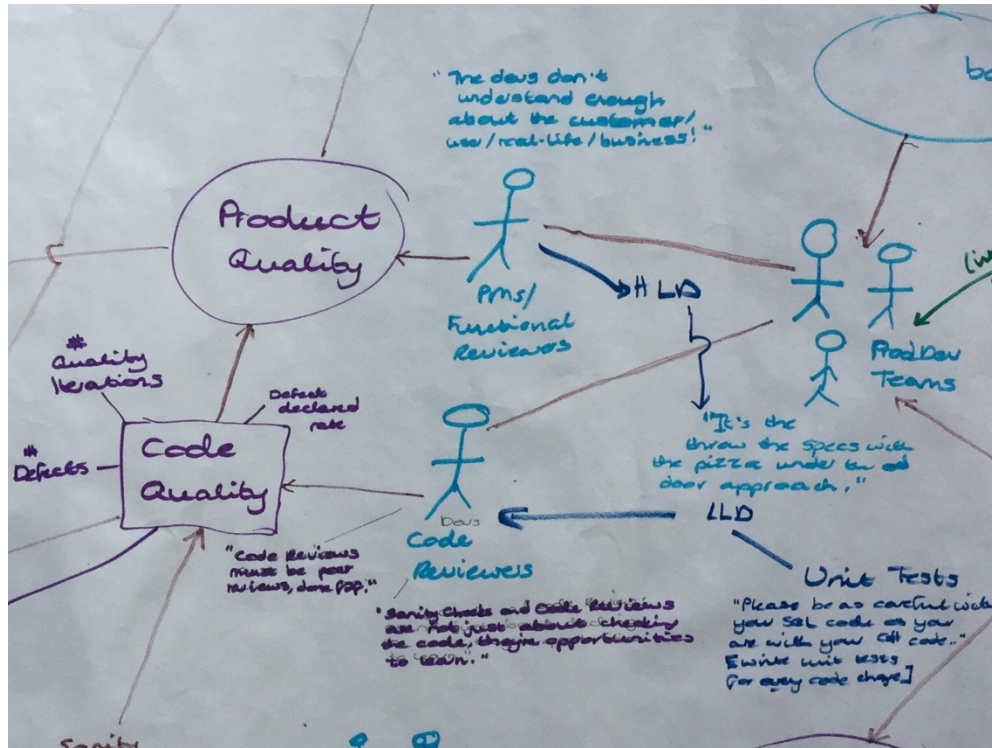


Figure 2. Richer picture excerpt, hand drawn.

A couple of weeks later, the authors re-drew the rich picture as a solely hand-drawn diagram. Even the small Post-It notes proved to be too large, and they didn't allow for as much flexibility or creativity as we wanted. Adding more connecting lines, colors, some drawings, free-form shapes and labels helped us to build richness into the diagram.

In this version of the rich picture, there are more characters: developers who review code for correctness, product managers who manage the requirements of the product being developed and development teams. High level and low-level design processes are now included, using the acronyms HLD and LLD. These acronyms are commonly used within the company and also save space on the picture. We have been able to group items into larger umbrella items, including Product Quality and Code Quality.We have added quotes to the rich picture, representing the sort of attitudes and beliefs that are held by people in various roles.

The tension around quality and productivity can be seen in the following quotes.

"Code Reviews must be peer reviews done face to face"

is commenting on the tension between the effectiveness of code reviews as learning experiences for improving code quality in tension with the time taken to do reviews.

> "Sanity Checks and Code Reviews are not just about checking the code, they are opportunities to learn."

This especially relates to new developers.

Near the bottom of this diagram, and the stick figure labelled Code Reviewers, and linked to Unit Tests there is something that the first author heard one developer say to another during a code review; explaining that TestFirst should be applied to every sort of code change, they went on, "Please be as careful with your SQL code as you are with your C# code!" (SQL code is used for accessing databases, whil C# code is used for implementing the function of the system).

For a business based largely around very analytical software developers, spending most of their days writing code, they rely on talking to each other. This is especially true around the issues of quality and time.

## An Even Richer Rich Picture

In an extensive open-plan environment dominated by large monitors and powerful desktops, there was very little paper around.

Our rich picture was on a large piece of paper, about A2- size. Because she wanted to keep it in sight and in mind, the first author left it laid out on the empty desk next to hers, for several weeks.

This provoked developers who came over to her desk to talk to her, and developers who were just passing, to comment on it and ask questions about it. It gave her unexpected opportunities to discuss what it represented and meant with the developers. Our understanding and interpretation of their work was made obvious to them, in a way that written text in a report, that they probably wouldn't read, would not have. It gave them a way to directly engage with and contribute to our fieldwork. Further, it gave us a unique way to validate our understanding of their situation with them, while the first author was there full-time.

This excerpt is the same section of the rich picture as that in the previous slide, but it is from several weeks later. It is a richer picture, in that it has had a lot of extra things added to it: more quotes, more interconnecting lines, more processes and text. Notably, the interconnections with design, side effects and product quality.

"It's the throw the specs with the pizza under the door approach", commenting on the concerns of the relationship between product managers and developers in what happens between the HLD (for which the product managers are primarily responsible) and the LLD (for which the developers are primarily responsible).

> "The devs don't understand enough about the customer/user/real-life business!",

relating to the concerns of the product managers.

Figure 3. Even richer picture excerpt.

Code quality is now embellished, and the associated processes are acknowledged by comments such as, "Having these policies makes me write better code".

On Quality Iterations, "It's just not enough to count the number of iterations … we need to know why it is happening." How many iterations (loops, occurrences) it takes to improve quality to an acceptable standard is not useful without understanding why it is happening. The quality iteration count is fast, but not necessarily useful to improving quality.

Over the next month or so, as we kept adding to the rich picture, our ethnographic understanding of the situation and the developers' software quality practices continued to deepen.

## The Complete Rich Picture

Below is the whole rich picture as it was at the end of the fieldwork period.

The phenomena of Software Quality, Productivity, People, Processes and Practices that emerged from the fieldwork are highlighted. as well as the overarching layers of communications, education, and Time.

What emerged was the components and people of the company demonstrably in rich, dynamic relationships.

Figure 4. Complete rich picture.

## TOWARDS ENTANGLEMENTS

A number of patterns, themes and connections emerged from analysis of the rich picture and the thick data that it represented. We would be writing about productivity and find that we were also talking about quality, and also people and time. It seemed impossible to talk about these elements separately. Looking for terms and ideas to express the strong bonds represented in the rich picture led to discovering previous work on entanglement, in particular Scott & Orlikowski (2014) and their use of Barad's (2007) notion of entanglement.

Scott and Orlikowski (2014)'s approach, which is based on Barad's agential realism theory of knowledge and being, gave a legitimacy to, as well as a way of articulating, the entanglements that emerged from our study. Scott and Orlikowski (2014) define entanglement as ``the inseparability of meaning and matter.'' These authors cite Barad (2007, p.ix), who explains,

> ``To be entangled is not simply to be intertwined with another, as in the joining of separate entities, but *to lack an independent, self-contained existence…* '' (our emphasis).

Barad (2007, p.ix) continues, "... Existence is not an individual affair. Individuals do not pre-exist their interactions; rather, individuals emerge through and as part of their entangled intra-relating." These individuals are not necessarily humans, but include non-humans, objects or phenomena involved in the situation we are trying to understand. Each of these is

not a discrete factor in reality, an independent object with "independently determinate boundaries and properties" (Barad 2007, p.33).

Barad (2007) describes these things as "phenomena" and sees them as relational, with their agency residing in that relating, rather than agency as something that resides in an individual thing. This is similar to Suchman (2007)'s understanding of agency. Barad goes further: when defining phenomena, firstly, in referring to phenomena as agencies, and secondly, and most significantly, that their existence and properties arise through their intra-acting with one another.

*Intra-action* differs from the notion of 'interaction'. 'Interaction' assumes that there are independent objects, or phenomena, each with their own agency, that precede or pre-exist their interaction or relating. Intra-action, however, is the mutual constitution of entangled phenomena: these phenomena come into being through their intra-actions.

Barad (2007) considers phenomena and their continual intra-actions to be constitutive of reality. Entanglements are dynamic, they are already made, as well as always in the making (Suchman 2012).

We realized that entanglement meant, in the first instance, that any attempt to understand the company in terms less than the whole rich picture, its elements and interactions, would lead to the understanding of a different company. And, in fact, a fictitious company.

A shift in our interpretation and representation of the local software development endeavor occurred in the move from initially exploring software development as a human endeavor, and as situated action (Suchman 2007), to a post-human perspective of entanglements in the local context. In the latter, humans and non-humans, their intra-actions and agencies, are seen as being equal participants, active in the ongoing, dynamic entanglements from which phenomena such as quality, productivity and practice come into recognizable being.

Viewing local software development as relating phenomena, and exploring the nuances of their intra-actions, makes entanglement a meaningful way of discussing the reality of software development practice. The entanglements of people's actions with phenomena such as quality, productivity and time, is a characteristic of the perpetually generated context in which the design and development of complex software is accomplished.

## "Slower today, faster tomorrow"

"Slower today, faster tomorrow" is one of the company's software development mantras. Experienced developers talk frequently about what this mantra means: if developers spend time and effort on assuring quality in their original code, then all of the developers will be more productive in the longer term. In other words, they will spend most of their time adding new functionality to the codebase, rather than spending time fixing defects that have been discovered in previously released (deployed) code.

> "So previously I would quite often talk about quality in the context of the speed quality trade-off… Because having quality gives you speed. So slower today and faster forever. So I've really toned back on my attempt to be fast and I've really just thought about how we can have quality instead. Because I don't even need to think about speed, I just get it automatically. So, for me quality is the ability for what we do now to have long lasting positive outcomes on the goals that we're trying to

achieve. So if we produce something that may take a little bit of time but in the long run saves us a lot of time then that was the right thing to do that, it's good quality."

Code that is not written well, that does not adhere to the company's coding standards, for example, is difficult to maintain and change later, and this in turn may lead to further defects and decreased productivity:

"If you don't write code in a good way, developers will spend more time reading and changing it, which will result in more waste at the end. It's all about our future speed."
"Particularly, I'm a software developer, so the quality for developers means we should write very elegant code. So, probably, for example, if we write, if I write, very dodgy code, there's a high possibility that my code would break something of the software or [worse] result in an unhappy client. Then they will lodge another incident and more repetitive work. So yeah, that quality [coughs] means, for me, is more work, more time—yeah, less productivity as well."

Increased defects in the code means that at some stage, the software will not work as expected, or worse, will crash while the customer is using it. Developer time will then need to be spent on fixing those defects, rather than spending that time on developing – and delivering – new features in the software.

"I mean, when we say we should deliver good qualities, there's always another thing called time frame. To deliver the good quality software, definitely we need more time. But normally people at [the company] got overloaded easily because if we got too much work, unfortunately we got too much defect as well."

The above quote highlights the tension between a stated value of spending time on quality, and the experience of time being scarce. However, spending substantial time taking action to improve the quality of the code is potentially detrimental to throughput and thus productivity.

After a developer at a daily team stand-up meeting said, in an ironic tone,

"Slower today, faster tomorrow!"

one comedian from the Productivity team responded, "but tomorrow never comes!"

They were reminding the team that one can spend forever getting something closer to perfect, or 'high quality', but, taken to the extreme, the work will never be delivered. This concern about not delivering 'enough' is not often explicitly articulated, but it is alluded to frequently and underlies much of the developers' everyday practices, behavior and decisions.

The issues of software quality and productivity in practice are about people's practices in time and over time. Decisions that the developers continually have to make include: what should we spend time on? how much time should we spend on what kind of work? should we spend more time on this work for better quality? if our throughput is higher in the short-term are we more productive in the long-term? how and where are people spending their time? and so on.

"Because I find a lot of the time when something goes wrong it's because - not that someone just did something silly, it's often that we didn't consider something. That if we thought about it for maybe half an hour longer, we could have."

This apparently simple, short phrase "Slower today, faster tomorrow!", frequently quoted by developers in discussions about quality, is really about the ongoing entanglements of the phenomena of quality, productivity, people's (developers') practices and time. This is illustrated by experienced developers' quotes above from their discussions about this mantra and what 'quality' means at the company. Moreover, it signifies how these phenomena are mutually constitutive: dynamically forming and shaping each other through their continual intra-actions.

## Developers Becoming

A "fully-fledged developer" a (human) developer comes into being through ongoing intra-actions with quality, productivity and technical development principles, processes and tools, and with the other developers, over a considerable time. These continual intra-actions generate entanglements within the local development environment and over months, the novice becomes a developer, and over years, they become a fluent, proficient developer. But they are not simply skillful developers; these developers are experts in the entanglements that are particular to the local environment in the participant company.

Producing high-quality enterprise software requires fluent, expert software developers, who have excellent programming skills, as well as the high-level technical skills to work with the automated testing system, sophisticated technology stack and other technological infrastructure used to continually build a complex, but robust, software product such as Connect. A reasonable amount of domain understanding of the logistics industry is also necessary in order to be able to work as an effective developer in this company.

The production of high-quality software requires new hires (developers) to gain both technical competence and fluency in the local codebase; both of these take time. The problem is not simply a concern that is regularly raised by more experienced developers, i.e., that new hires lack the necessary technical skills and expertise to be productive and produce quality code, i.e., code that is maintainable, efficient and thoroughly tested. It is also about the continual trade-off for senior, experienced developers between mentoring, or coaching, of new developers, which takes considerable time, and getting their own development work done in a timely manner:

"All the senior devs. are already busy doing what they are working on at the moment. It's kind of like, they've got their work and they have to teach other people at the same time. So the priority for senior devs. is, of course, their current work."

"Yeah, and it takes a lot of time as well. Sometimes my manager asks me to be a mentor to the new developers, but I'm already overloaded and then this new people come and ask me, 'How can I do this? How can I do that?'. Sometimes it's really annoying. If I didn't have enough work to do, I'd be more than happy to help them, but the reality is not like that."

These quotes from senior developers in two different teams make the point that the senior developers' most important focus is their 'current', i.e.,technical work, their design and development work, and this is what they 'should' make their priority, in order to be productive. Mentoring newer developers is an extra, 'really annoying' impost on their time and effort. They do not view mentoring – getting newer developers au fait with the company's development systems, processes and tools – as being as valuable a use of their time and expertise as producing software themselves.

The next quote from a technical team lead refers in part to the assumption that it takes a certain amount of time for any developer to become expert enough to both produce quality code themselves, but also to make assessment of the level of quality of another developer's code:

> "A typical situation, when some developer jumps from junior level to let's say senior level his complexity of work rises, it's natural that number of defects can grow as well but it's kind of natural at first… but I'm going to introduce it and what I'm going to do, I'm going to assign that task to junior developer capability. They need to learn how to do code review because it's - a typical situation… I'm not ready to give them proper final code reviews but at least I think if I give them these intermediate code reviews maybe they can improve their code in quality as well. Because typically it's, I don't know, sometimes it's as long as two years for a developer to gain my trust, so I progress the developer to a capability which allows code reviews."

The aim for a developer's performance is that they become fluent in producing complex code in a collaborative development environment. The more fluent a developer is, the faster they will produce code. And, crucially, they will not only code faster (than a developer who is not as fluent), the code that they will produce will be a higher quality code, without requiring as much iteration or revision. They are therefore more productive as individual developers. Further, they will be able to do code reviews of other developers' code more effectively, which will improve that code's quality. And, if this developer is mentoring a new developer, the less experienced developer will be coached to write higher quality code. So, a secondary effect that is hoped for is that the reviewed developer's approach will change, or at least shift, so that the code they write in future improves also. This impacts productivity in two ways: firstly, it ensures that the particular piece of code in question that is checked in to the code base and eventually released to customers is higher quality, and secondly, fluency of the newer developer improves which, in turn, will greatly improve the code that they produce in daily practice. This will then reduce the need for iterative code reviews at the development end and/or defect fixing at the production end.

## Insights from entanglements

The entanglements that are central to our understanding of the local software development situation are those arising from the intra-actions of quality, productivity, people, practices and time. They are not the only ones in the local situation, of course, but these are the ones that emerged most persuasively from our fieldwork and analysis. A researcher's observations in any situation are always limited in various ways, and we can therefore only ever have partial knowledge of it (Haraway 2001).

Perceived software quality and productivity levels unfold as a result of the ongoing intra-actions over time of the developers, their everyday practices, company software quality and productivity principles and processes, development infrastructure and other undefined (in this fieldwork and study) phenomena. Ultimately, levels of quality and productivity in the company depend entirely on the developers' everyday actions that make up their practices. In the end, it is only what the developers do – the actions they perform day-after-day, over long periods of time – that matters.  It is the intra-actions of practices (actions), quality, productivity and time as developers continually attempt to balance the demands of quality and productivity, and the efforts given to achieve one or the other, or both, over time that give rise to ongoing entanglements. These entanglements mutually and simultaneously form these phenomena. The phenomena that we identify as practices, quality and productivity are becoming; they continually come into recognizable being through their dynamic entanglements with each other, time and the developers themselves.

These entanglements give us some insight into the subtle complexities of this kind of software development work and the expertise and technical fluency required to carry it out effectively. They also give us a way to describe the continually generated context in which the collaborative design and development of complex software is accomplished.

Somewhat ironically, taking a human-centric stance led us to conclude that quality and productivity in software development requires more than simply focusing on the humans (software developers, in this case).  Applying Scott and Orlikowski (2014)'s Baradian approach to reality as ongoing intra-actions of phenomena gave a legitimacy to, and a way of articulating, the dynamic entanglements that emerged from our study. Recognition of these entanglements shifted our perspective from a humanist one, focused on collaborative software development as essentially a human endeavor, to a post-human appreciation of the setting's complexities and the mutual constitution of the phenomena central to our research focus, i.e., the developers and their practices, software quality, productivity and time.

## LESSONS LEARNT AND ORGANIZATIONAL IMPACT

The concept of entanglement provides an explanation of the local situation as dynamic, multiple and emergent.  Together with the rich picture, it

- presents the nuances of the developers' everyday work practices as they are constituted within the local situation; and
- builds trust with participants, as they see an attempt to capture and express the complexities of software development and their lived experience of it.

This research had a significant impact on the organization and our continuing relationship with it and the developers.

By making our ethnographic work visible through the rich picture, and encouraging participant developers to make suggestions or additions, there is a sense in which they jointly own this work.

The rich picture continues to evolve, and is now explicitly owned and edited by the organization, and used to explore software quality concerns, with our oversight.

The ethnographic insights that we shared with the participants helped us to secure support for, and engagement with, subsequent experiments in mentoring and measurement. The aim is to help them develop practices that will sustain, even increase, software quality, in the face of particular challenges. These are the continual growth in the size, complexity and customer reach of the Connect codebase, and the ongoing hiring of new developers unfamiliar with the organization's quality principles and practices. The work will be characterized by participatory methods and deep collaboration with the developers, enhancing the potential future organizational impacts.

**Julia Prior** is an Associate Professor in software engineering at UTS. She is a software developer, an ethnographer and a teacher. Her research focuses on understanding the lived experience of professionals developing large, complex software systems and the mechanisms that enable effective collaboration and quality assurance. You can contact her on <julia.prior@uts.edu.au>

**John Leaney** is an Adjunct Professor in software engineering at UTS. Over the last fifteen years, he has developed expertise in combining qualitative techniques, such as action research and ethnography, with quantitative approaches to provide effective methods for understanding and designing architecture-focussed, complex software systems.

## NOTES

## REFERENCES

Barad, K. 2007. Meeting the universe halfway: Quantum physics and the entanglement of matter and meaning. Duke University Press.

Checkland, Peter. 2000. "Soft systems methodology: a thirty year retrospective." Systems Research and Behavioral Science: 17, S11– S58.

Haraway, Donna. 1988. "Situated Knowledges: The Science Question in Feminism and the Privilege of Partial Perspective." Feminist Studies, 14(3), 575-599.

Mistrik, Ivan; Richard M. Soley; Nour Ali; John Grundy; Bedir Tekinerdogan, eds. 2016. Software quality assurance: in large scale and complex software-intensive systems. Morgan Kaufmann.

Prior, Julia. 2011. Everyday practices of agile software developers. PhD dissertation, University of Technology Sydney.

Scott, S and Orlikowski, W. 2014. "Entanglements in practice: Performing anonymity through social media." MIS Quarterly, 38, 873–893.

Suchman, Lucy. 2007. "Human-machine reconfigurations: Plans and situated actions."  Cambridge University Press.

Suchman, Lucy. 2012.  "Configuration." Inventive methods, pp. 62-74. Routledge.

# Ghost in the Machine

## How Taxonomic Metadata Allows for Scaling Ethnographic Insights into Search Algorithms

AMANDA KRAUSS, *Duo Security (formerly Indeed)*
ALEXANDRA TEODORESCU, *Indeed*
LEORA YARDENAY, *Indeed*

*This case study explores how we personalized search results by turning ethnographic insights into taxonomic metadata, which in turn allowed us to use quantitative methods to assess business impact. The first part of the case study focuses on the problem we were trying to solve—creating better search results for nurses—and using ethnographic interviews to understand how nurses approached looking for jobs. The second part of the study dives more deeply into how metadata works, and why it was the perfect partner for capturing our ethnographic findings and making them into a scalable and measurable part of the design process. The third part of the study details how we tested and scaled our designs in the live project, and why we believe others might benefit from using a similar approach. Keywords: mental models, taxonomy, business impact.*

## INTRODUCTION

While researchers understand how valuable ethnographic research is for defining a problem space—particularly in the early stages of product development—it can sometimes be difficult to justify the time and cost of this method to employers and stakeholders. All too frequently, companies focus on "quick wins" using remote usability testing or A/B testing, the better to show impact with quantitative metrics. Furthermore, it can be challenging to articulate the value of discovery research within this environment, particularly because the perceived value of the research often ends with the report. Given this setup, it is difficult to trace a direct line of impact between ethnographic research and large-scale, measurable results.

In this case study, we show how combining both qualitative and quantitative approaches—specifically, ethnographic research, taxonomic metadata classification, and product analysis—can measurably improve search engine results. We begin by describing the ethnographic research that allowed us to understand how nurses searched for jobs. We then detail how we used taxonomic metadata as a vehicle for our findings, as well as describing the value of metadata for scaling the project. We outline how metadata is a tool for disseminating users' mental models throughout the site; how it can inform both back-end (algorithmic) and front-end (user interface) design; and how, once the metadata was in place, we were able to repeat the process with other segments beyond nursing and in markets outside the U.S. Finally, we will demonstrate how this approach addressed the aforementioned challenges in measuring the impact of ethnographic research, and allowed us to draw a direct line from our initial discovery interviews through to the A/B testing of the final design.

By sharing our process and outcomes, we hope that this case study will help other ethnographic researchers find ways to think about measuring the impact of qualitative research, as well as strategies for scaling the results of qualitative insights.

## BACKGROUND

Indeed.com is a global job search website that receives over 250,000,000 visitors every month.[1] The core functionality of the site is the search engine: employers post jobs, and job seekers look for them. As researchers, designers, and taxonomists, our goal is to provide the best search experience possible, and to strive for continuous improvement. In 2018, one major area slated for improvement came from the realization that, at its inception, our site had been designed for an "average" job seeker and that—based on many teams' research—there likely wasn't an "average" job seeker, but rather, different segments of job seekers with different wants and needs.

With the goal of more personalized search results on our minds, both the User Experience (UX) and Taxonomy teams were trying to facilitate cross-team work. One organizational challenge was a presumed separation of interests: UX was tasked with designing the front-facing user interface while teams such as Taxonomy were tasked with the designing the back-end of the site—meaning the algorithms and the metadata that power the search—without having established avenues to test the potential front-end design impact of their work.

In 2018 the Taxonomy team had created an extensive metadata system for classifying jobs in Occupation categories. In layman's terms, Occupations are groups of jobs with similar duties or responsibilities, such as nurses or truck drivers. The design team knew that this data could be used to make design decisions, and the Taxonomy team understood the value of getting user feedback on their work. It was an opportune time, then, for the creation of a new, cross-functional team that aimed to bring together both qualitative and quantitative analysis to ensure the best user experience for specific segments of customers, while also tracking business impact. The team was called "Segmentation" and its overarching goal was to examine and improve our business results for certain segments of job seekers. A segment could be defined in many ways: it could be a category (such as people who work from home); an industry vertical (such as government work); an occupation (such as lawyers or truck drivers); or a behavior (stay-at-home caregivers). Regardless of classification, the team's director believed strongly in the need for ethnographic research as a key tool for understanding our users.

## RESEARCH GOALS

Based on market research, the first customer segment identified for our project was nursing. We already knew that our nursing job listings were not performing as expected, and according to the Bureau of Labor Statistics, the market for nursing jobs was only projected to grow (U.S. Department of Labor 2020). And so, starting the segmentation project in August of 2018, we wanted a deep, foundational understanding of what job searching looked like for nurses. As the project progressed, our objective became more focused: based on what we'd heard, we needed to ensure that the search engine accurately reflected the mental models of our nursing job seekers; to do this we needed to iterate and validate both

taxonomy choices and user interface designs that surfaced new metadata for feedback. In the final phase, we had to find ways to measure the effect of these design and taxonomy decisions by running in-product tests to show business impact.

## METHODOLOGIES AND PROJECT STRUCTURE

In the course of the project, we used several methodologies, ranging from ethnography and card sorting to data analytics and A/B testing. For this reason, it is easiest to walk chronologically through the project, examining each method as we go.

### Taxonomy-Informed Ethnographic Interviews

In order to collect user feedback from nurses, we needed to find the right participants for the interviews. The Taxonomy team had definitions in place for who was or was not a nurse and had recently created search facets specific to nurses, such as licensing credentials. We based the interview screening process on the existing criteria and filtered out potential participants that didn't fit the study profile. For example, we ruled out Physician Assistants, Surgical Technicians, and Phlebotomists, but included Registered Nurses, Licensed Practical Nurses, Licensed Vocational Nurses, Nurse Practitioners, as well as Nurse Assistants and Aides. We then interviewed fifteen nurses, using a semi-structured approach that allowed us to ask about job search habits, desired jobs, and other areas of inquiry.

From our ethnographic research, we found that nurses, too, had formed opinions about nursing as a job category. As one nurse told us, "I think that there are certain things that shouldn't even be in this category. If I'm searching for an RN [registered nurse] job, medication aide doesn't need to be there, or nursing assistant, or home health aide. A nurse assistant aide is not going to look for an RN job. To be honest, that's a little offensive." This information bolstered our suspicions that we were on the right track with trying to categorize jobs in a more refined manner.

When it came to searching for jobs, we also found that nurses had a great deal of information at their disposal. Due to their deep professional networks, they usually understood the options for employment in a given locale, and also had a good idea of how well local employers paid. Additionally, experienced nurses had options, and therefore strong preferences, about how and where they worked. One nurse saw a strong correlation between the appeal of different types of nursing jobs and specific personality types: "The Emergency Department can be kind of blunt: very concise, quick, not super in depth—Type B nurses. ICU [Intensive Care Unit] nurses are the flipside. They are very intensive with their care. They are very Type A. Medical Surgical is the 'dark and stormy place' of nursing. Most people don't necessarily want to work Med Surg, but it's known as a good place to start out."

Because experienced nurses had very specific requirements, they were frustrated that the search interface, including the job descriptions themselves, did not include information addressing these preferences. They wanted more information about the working environment (including equipment), what types of patients they would be working with, and detailed information about shifts; unlike typical office jobs, nursing can include anything from four- to twelve-hour shifts. Finally, they needed information on which medical specialties were expected in a given job. Overall, our analysis revealed what Indi Young calls mental models (Young 2008). In our case, the nurses' mental models were the preferences

and information needs underlying their decision-making process. As Young notes, identifying mismatches between your product and customer expectations can provide opportunities for designing better solutions. Based on our interviews, we had identified clear gaps in our product; the challenge was to address these gaps within the framework of a mature and complex search engine. We could not rely on our generative research being integrated into initial product development, nor was it feasible to add surface-level features on top of the existing product. Furthermore, an individual job description was unlikely to provide the information nurses wanted. For these reasons, we began to think about how to leverage metadata to solve our problem.

## Method in Detail: Metadata as Vehicle

Before talking about the next step of the project, it may be beneficial to talk more generally about the role of taxonomic metadata in search-based products. One example is the system an e-commerce site uses to classify its items. A shoe site might classify shoes using categories such as boots, slippers, types of heel, or color. Similarly, a clothing site might use descriptive taxonomy categories such as "billowy" or "bohemian" (McDowell 2020).

All of these descriptive items are taxonomic metadata. The data may have been created manually by humans, just as one might apply category or tag labels to articles in a content management system; they may also be created by running items through various types of automated analysis algorithms; finally, and as we will discuss below, taxonomic metadata is often the result of combining manual classification with automated analysis. Importantly, while a site's user interface features (such as filters) may expose the taxonomic metadata directly to users, the metadata can also exist as an invisible support system that helps turn user queries into search results—a function most often associated with query understanding. Any type of metadata can be made visible or invisible to the end user, depending on how a given platform chooses to display it (Gartner 2020; Harvard Law School 2020).

Indeed uses a search engine just as an e-commerce site does, but instead of searching for clothing items, our users are looking for jobs. The content of Indeed's search engine is made up of documents related to job searches; here specifically we will focus on job posting documents which employers create and upload, either to our site or other online locations. The job documents are then indexed by Indeed's search engine, allowing job seekers to find them using search queries. When employers post jobs, they may not explicitly include all information the job seeker wants. For example, they may post a job with the heading "Project Manager," a title which lacks relevant industry context: this job could exist in areas as diverse as construction, electrical work, or software engineering. By analyzing job titles in conjunction with information contained in job descriptions, the Taxonomy team can create relevant Occupation categorizations, and add this metadata to the job document.

Once the metadata is added, it can remain hidden from view, and be used as a factor in ranking search results—this is what we might call a "back-end" use. Alternatively, its presence can be made visible with design choices on the front end, such as filters. The value of metadata, then, is that rather than being a single feature, it has the potential to power many different features. It was in this context that we began to think seriously about how to translate the findings from our ethnographic discovery research into metadata usable by the search engine, and by extension, able to be disseminated into the site at large.

## Combining Design and Analysis for User-Informed Taxonomy Development

As of 2018, Indeed's Taxonomy team had done extensive work on developing its own metadata classification schemas, including a home-grown Occupation taxonomy. It is important to note that while there are existing third-party categorizations of jobs by industry and occupation, like the Bureau of Labor Statistics (BLS), they do not update categories regularly; it might well be a decade between updates. For this reason, the BLS does not reflect the quickly changing nature of jobs in the modern world, and also falls behind on linguistic trends (e.g. the switch from "Library Science" to "Information Science"). Because Indeed's core content is centered on what job seekers want, the Taxonomy team must be able to move in a more agile fashion, and address search queries happening in the moment. A recent example of this need was the advent of COVID-19; based on the work we are describing here, Indeed's teams were able to incorporate this data into medical job listings quickly. Last but not least, our Taxonomies are created with a global audience in mind.

In order to create taxonomic metadata, the Indeed Taxonomy team uses a variety of techniques, from external research and qualitative studies (e.g. card sort), to quantitative analysis of internal data and text mining. Creating the classification schemas, however, is only one aspect of metadata at Indeed. We also need to extract the metadata concepts we define from unstructured text, like jobs and resumes, so they become available throughout our infrastructure and can be used in a variety of search products and algorithms. We do this via a combination of natural language processing (NLP) tools and human curation.
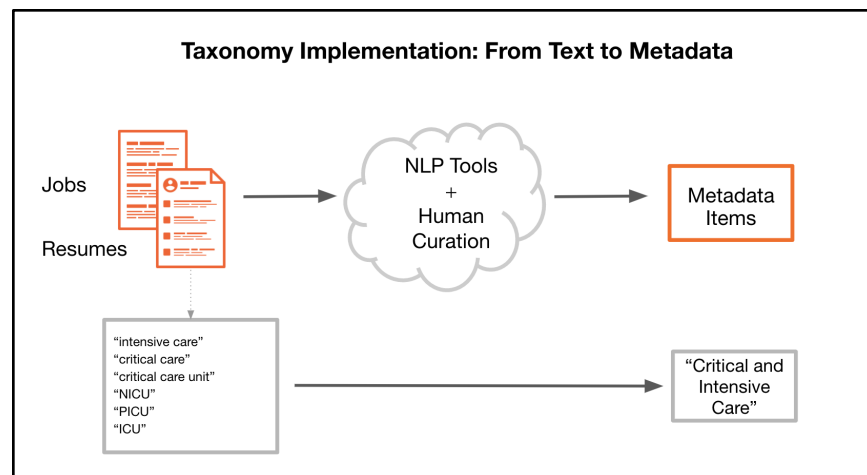


Figure 1: Flowchart showing how the Taxonomy team maps text to metadata.

In short, the taxonomic metadata captures information that job seekers expect to be relevant—that is to say, it expresses something of their mental model in ways that the original document might not. In this way, metadata can improve the search engine's performance so that users find the best results and the employers find the best candidates. Ultimately, our metadata helps bridge the gap between what users intend when they type something in the search box, and what the site returns to them.

For this reason, the Taxonomy team had already considered using medical specialty metadata as a search facet for healthcare Occupations, as early as 2017. But given product development tradeoffs, the team had to prioritize other work. The ethnographic research with nurses provided a crucial opportunity to revisit that priority. The interviews revealed just how important the information was to nurses, which suggested that it might also be important for other medical job seekers. Additionally, the Segmentation team's market research suggested that it was an opportune time to move medical taxonomy work from the back burner to the top priorities list. Given the opportunity to work more on medical specialties, the Taxonomy team had to make the choice between adopting existing taxonomies of medical specialties, like the one created by the American Board of Medical Specialties, and building a new taxonomy that would fit our job seekers' needs and the data trends on our site. The Taxonomy team also had to make internal decisions on whether they wanted separate taxonomies for medical specialties and medical departments, as opposed to merging them into one category, and on whether medical specialties should be specific to nurses or cover all health care professions. And in order to get answers to these questions, the team needed to surface the taxonomy categories directly to users via the front-end design. In short, the ethnographic research created another opportunity: a chance to expand on the existing taxonomies in a user-informed fashion.

## Design Testing and Iterating on Taxonomies

In January 2019 we began our first design tests, in which we tried surfacing skills metadata as a filter in a design concept (as illustrated below).
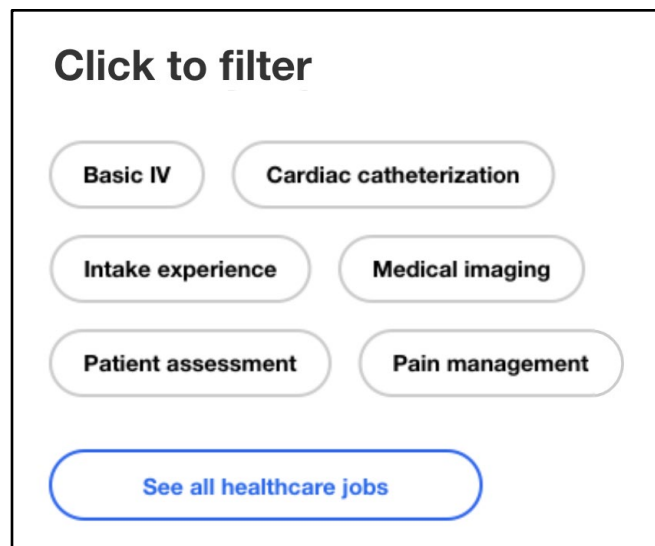


Figure 2: Design for the initial concept test offering skills-based filters such as "patient assessment" and "Basic IV."

Our research sessions revealed that filters were a successful approach, but our first attempt at surfacing data didn't quite hit the mark. As one nurse noted: "I guess I'm a little

confused about the 'Basic IV' because that's required by all nurses. So I think that's kind of redundant." As non-experts, we had made the wrong call about which pieces of data it was important to display, and it was crucial to understand that early on. And since the Taxonomy team was working on medical specialties in parall el, this feedback helped them decide what the final taxonomy categories were going to look like.

As work progressed in 2019, the Taxonomy team was able to quickly update and refine categories to ensure they fit our users' mental models by closely collaborating with Segmentation and UX Research. Unlike the skill-based filters we'd started with, the medical specialty filters (for example "medical imaging" or "cardiac catheterization") performed well in our initial design testing.

**In-Product Testing and Agile Taxonomy Feedback**

Having already done small-scale design testing with the filters, we were ready to expand our testing to in-product analysis. In the third quarter of 2019, the Segment Manager launched an A/B test using the ethnography- and taxonomy-informed filters we had been designing. A/B testing allows a site to show two different versions of a user interface to two different sets of users; logs of user behavior are used to identify which group (A or B) performs best, according to the metrics that have been chosen. As mentioned earlier, Indeed has a large number of users so we can run A/B tests quickly and easily identify statistically significant results. Based on the user behavior, the product stakeholders decide whether to modify the product based on new designs or to maintain its current status.

In this case, we looked at analytics to identify the top nursing related queries and targeted these users for the A/B test. 50% of the users (the control group) experienced the normal interface on Indeed, and 50% were in the test group, where they saw the medical specialty filters. In the first study, our test group added three filters to the site: Specialty, License, and Patient Type.
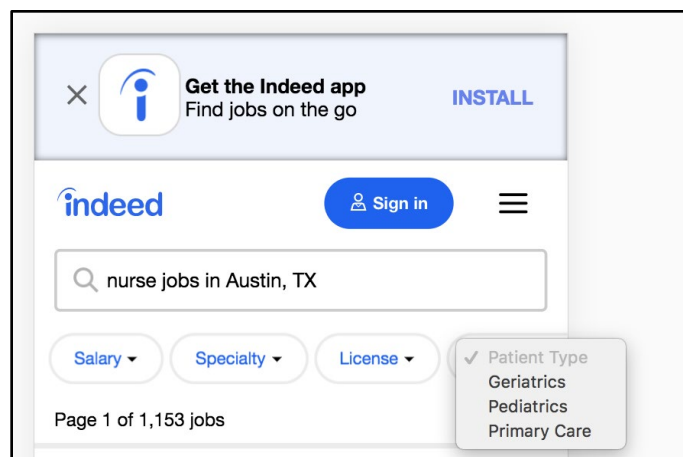


Figure 3: Design for the A/B test of nursing metadata, using mobile filters created by ethnographic research combined with taxonomic analysis.

For the Patient Type filter, we started with three categories: pediatrics, geriatrics, and primary care; note that these categories and options were prioritized based on what we'd heard speaking to nurses in our ethnographic interviews.

Unfortunately, this filter did not perform well in initial tests. After further investigation and discussions we decided that what we'd initially thought of as "patient type" fit better under the "medical specialty" data category. With further testing, we narrowed the design and taxonomy combination down to two filters, Specialty and License.
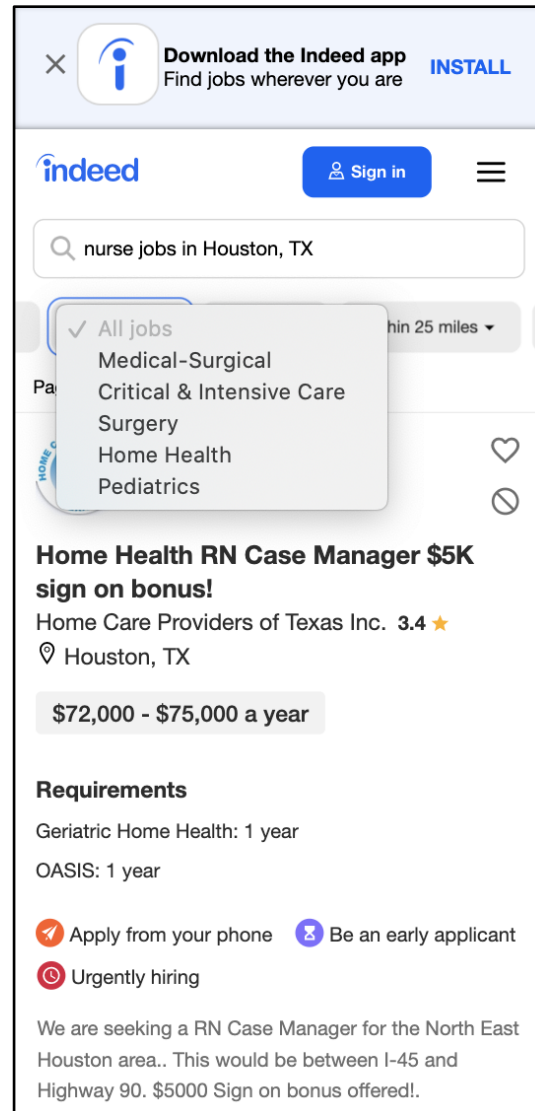


Figure 4: A finalized version of the mobile filter, which combined license and medical specialty.

Ghost in the Machine – Krauss et al.

Even after narrowing down the data itself based on A/B testing, the Taxonomy team ran further card sort studies to improve filter display names. The goal was to identify names of specialties that were familiar but also distinct enough: for example, to decide whether "OR" (Operating Room) was a better label than "Surgery." There was also a need to test the boundaries of conceptual overlap, and to find out how similar or dissimilar "medical-surgical" and "surgery" were for nurses.

## MEASURING AND SCALING THE RESULTS

Based on the A/B tests, the Segment Manager was able to produce a quantitative analysis of the results. Indeed uses multiple proprietary internal tools to measure A/B tests, and they include a homegrown version of Structured Query Language (SQL) which allows us to pull mass amounts of data from our databases. The Segment Manager created queries to see which filters were used most often, which were underutilized, and which resulted in an application to a job.

After monitoring weeks of data from thousands of users, overall filter usage increased and we also saw an increase in positive outcome metrics. These results aligned with the team's KPI for the quarter (although we cannot share details for legal reasons). Thus, by addressing the issues nurses reported in our initial ethnographic interviews, and supporting those needs via taxonomic data, we could show clear, positive business results.

Later, having seen success with the medical specialty and license filters, Segmentation decided to perform another A/B test using a "shift" filter. Taxonomy teams had been building these attributes because during the ethnographic interviews, the nurses had identified scheduling as a major pain point that influenced their decision to apply to jobs. Attributes included filter options such as shift length, time of day, and hourly.
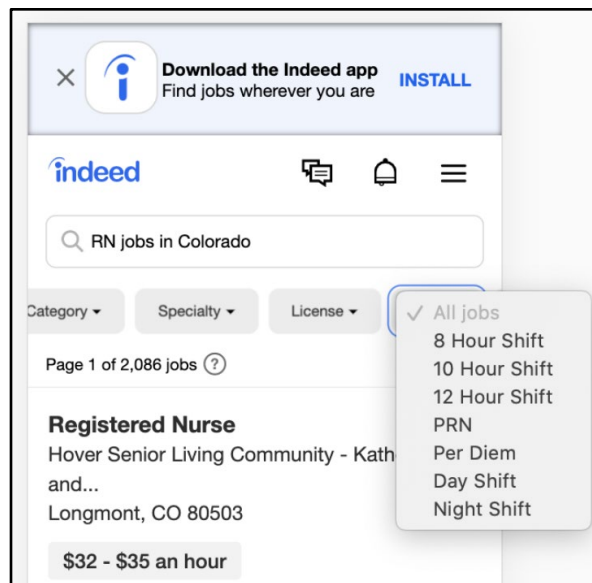


Figure 5: The nursing "shift" mobile filter, which displays multiple options to job seekers to account for their unique needs.

What happened next was interesting. After reviewing the results of the shift filter test there were no statistically significant changes in desired metrics. In the past, without any meaningful change, a team would turn off a test and deem it unsuccessful. Yet, rather than dismissing the data, Segmentation decided to create a new process. If results were neutral (or trending positive), they would roll out the changes in production. The logic behind this change was that customization is valuable in and of itself; both psychological (Devaney 2017) and ethnographic insights combined to tell us that customization and additional information was valuable. Job seekers want to feel that their unique needs are being heard and addressed, and therefore, a design that shows net neutral A/B metrics is still worthwhile as a way to personalize the site for them.

Furthermore, once disseminated throughout the site, the metadata continued to be used in designing experiences; as Figure 6 shows, the data is now included regularly on the job description.
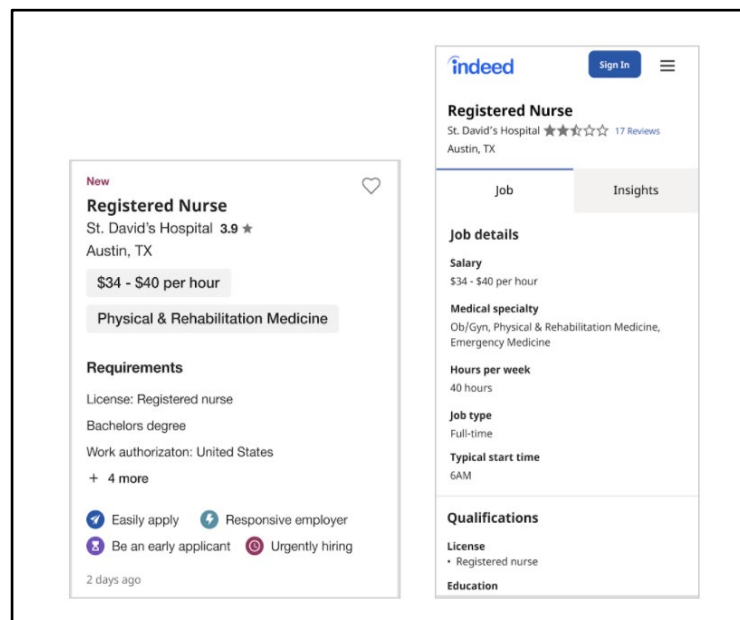


Figure 6: The "job card," which is a job description that displays structured data. In this case, the nursing metadata created by our analyses became a regular part of the job description display.

In this way, starting from our ethnographic understanding of how nurses make job decisions, we moved to large-scale implementation of the necessary information using taxonomy metadata as a vehicle. And this was only the first segment that we pursued; starting in 2018, we expanded the same approach to other segments such as trucking and retail. We now had a rough idea of how best to pursue the journey from initial discovery work to in-product testing, and were able to streamline the process.

Today, we continue to scale our work domestically and abroad. For instance, what started as four medical specialties is now fifty-seven medical specialty attributes in the USA. Shifts and schedule attributes now exist in approximately twenty-seven markets and requests are coming in regularly to expand our taxonomy to reflect international needs. By the time of publication additional attributes will have been created and may even be visible on our site.
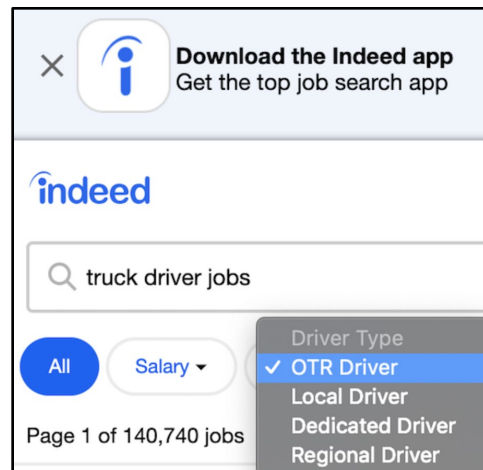


Figure 7: Trucking mobile filter for job type, using structured data generated by the same ethnographically-driven process as we used for nursing.

Working together across teams, we were able to combine qualitative and quantitative analyses in ways that benefitted both our internal and external users. By having ethnography inform taxonomy categories, and vice versa, we were able to create metadata that informs both the front-end and back-end of product development, and ultimately ensures that ethnographic research makes its way into the front-end and the back-end of the site.

**Amanda Krauss, Ph.D.** is a Senior Design Researcher at Duo Security (now a part of Cisco), and was previously a Senior UX Researcher at Indeed. She has been doing mixed-methods research for over 20 years, and has worked in digital media, academia, and consumer-facing product development. *akrauss@duo.com*

**Alexandra Teodorescu, Ph.D.** is a Taxonomy Analyst at Indeed where she develops metadata structures and ways to extract information from text. Before Indeed she taught and conducted research into Linguistics and Cognitive Science at The University of Texas at Austin, Princeton University, and McGill University. *ateodorescu@indeed.com*

**Leora Yardenay, M.Ed.** is a Segment Manager at Indeed, where she conducts qualitative research and product analyses. After graduating from Tulane University, she served with AmeriCorps and pursued a Master's in Education from The University of Texas at Austin. She then pivoted to tech and currently works at Indeed, the employment-related search engine. *leora@indeed.com*

## NOTES

1. This figure represents the number of unique visitors to the site according to an internal Google Analytics report run in September of 2018.

## REFERENCES CITED

U.S. Department of Labor, Bureau of Labor Statistics. 2020. Occupational Outlook Handbook. Registered Nurses. Accessed August 21, 2020. ttps://www.bls.gov/ooh/healthcare/registered-nurses.htm

Gartner. 2020. "Gartner Glossary" *s.v.* metadata.  Accessed August 21, 2020. https://www.gartner.com/en/information-technology/glossary/metadata

Harvard Law School. 2020. "What is Metadata?"  Harvard Law School website. Accessed August 21, 2020. https://hls.harvard.edu/dept/its/what-is-metadata/

McDowell, Maghan. 2020. "Taxonomy is the New Fashion Tech Essential." *Vogue Business* website. July 28. Accessed August 14, 2020. https://www.voguebusiness.com/technology/taxonomy-is-the-new-fashion-tech-essential-the-yes

Devaney, Erik. 2017. "The Psychology of Personalization: Why We Crave Customized Experiences." *Hubspot* website, July 28. Last modified April 17, 2017. Accessed August 21, 2020. https://blog.hubspot.com/marketing/psychology-personalization

Young, Indi. 2008. Mental Models: Aligning Design Strategy with Human Behavior. New York: Rosenfeld.

**THEMATIC SESSION**

# Scaling Research

In this session we'll play with scale in multiple senses. We'll look at value and evaluation, sample size and team size, and consider how the impact of ethnography scales in different ways.

Session Curators: Laura Cesafsky, Kat Ekberg, Evan Hanover, Scott Matter

PECHAKUCHA

# Researching the Researcher

## How I Used Data Analysis to Understand my Reading Habits

STEPHEN Ó MATHÚNA, *Workday*

*The digital world has given us unprecedented access to information about ourselves. As human beings we can quantify ourselves on the basis of how much we eat, how much we exercise, how many miles we've travelled, among dozens of other facets of our lives. As technology gives us access to more and more 1s and 0s, our ability to measure and codify ourselves grows exponentially.*

*But these data can tell us stories, if we take the opportunity to stop and reflect on them. In this talk I examine one facet of my life - my reading habits - and put it under the microscope. I try to learn more about myself by examining seven years' worth of raw data, collected across 200 books. Using the research findings I make recommendations to myself about my reading habits, addressing areas such as author diversity, genre variety, among others. In addition, I explore the power of books in evoking emotional memories. Books anchor us in place, in time, and in emotional contexts. I argue that these are invaluable opportunities, in a world that moves so fast we don't often have a chance at such reflection.*
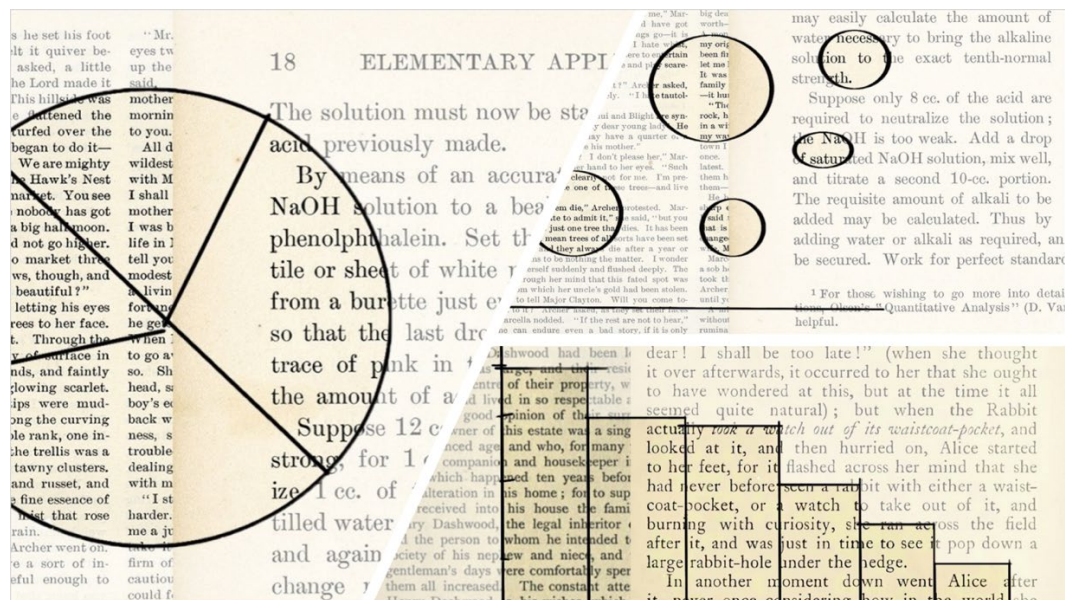


"The data behind the stories", © Stephen Ó Mathúna

**Stephen Ó Mathúna** is a UX Researcher based in Ireland. In 2015 he swapped academia for the world of product design and research. Since then he has applied his passion for research in the industries of augmented reality, education, travel, and human capital management.

CASE STUDY

# Beyond User Needs

## A Meaning-Oriented Approach to Recommender Systems

IVETA HAJDAKOVA, PhD., *Stripe Partners*
DEB MCDONALD, *Spotify*
SOHIT KAROL, PhD., *Spotify*

*This contribution is a case study of Spotify, a popular music streaming app, which uses automated recommendations to provide a better user experience to its listeners. Automated recommender systems have mostly been built around understanding user needs and user goals. Our case study presents a meaning-oriented approach aimed at understanding what users regard as meaningful and how an automated recommender system can forge meaning and offer experiences that help develop existing connections to music and generate new ones.*

*Following the meaning-oriented approach inspired by Lucien Karpik (2010), we were able to better understand how different audience segments engage with music and experience music as meaningful. We identified 2 cultural engagement models that listeners use to relate to music: (1) musical engagement during which music is the focus of the experience; and (2) non-musical engagement, during which the listener is the focus of the experience. Each engagement model uses different types of cognitive and evaluative aids, which we refer to as cues and proof points, to derive meaning from listening experiences. We also identified nine distinct types of experiences of meaning defined by distinct types of cues and proof points.*

*The proposed approach is applicable to the study and innovation of experience-led digital platforms and recommender systems.*

*Keywords: meaning, recommender systems, music, streaming*

Scale is a particularly urgent theme when researching and designing for digital platforms, algorithmic technologies and the attention economy. Some streaming platforms, such as Spotify, Netflix and YouTube, are based on business models which require them to acquire millions of users and provide value by creating customized, engaging experiences. In order to do that, these interfaces need to be automated so they can harness data to offer personalized content relevant to users' tastes, contexts and moods.

In order to provide delightful listening experiences in every session for every listener, Spotify faces specific challenges and opportunities related to the affordances of its main medium, sound. Spotify is one of the unique applications where most of the user experience happens through people's ears, brains and bodies, as opposed to their eyes. This poses several challenges when trying to understand what the value of user experience is and how to improve it.

First, unlike visual interfaces, where in-app interactions like time on the screen, likes and saves are the behaviors that could be used as proxies for understanding the value that users are deriving from the product, a lot of listening sessions on Spotify are hands and eyes free. This means that once users hit 'Play' and listen, we know very little about what experience they're having. On top of that, any behaviours with the visual interface are highly driven by

the context that listeners are in. For example, activities like running and driving a car, by their very nature, prevent people from interacting with the visual interface.

Second, based on the attitudinal segments, different types of users have different expertise in music and different abilities to navigate music, find music they like and discover new music. Their metaphors, expectations and benchmarks for deriving value from listening to music differ significantly and therefore, they require different forms and levels of support and feedback.

Thirdly, music listening itself is contextual. A playlist that is relevant at work might have a totally different meaning when listened to with kids at home and the measure of value evolves continuously with context.

Finally, while Spotify had a good understanding of what users' needs are in various contexts, scenarios and use cases, there was a gap in understanding what they experience as meaningful. Our challenge was to understand how and why users derive meaning from music and how we might train the recommendation algorithms to respect that nuance of human experience.

Therefore, the underlying research challenge was: how do we scale automated recommender systems to forge meaning and offer content that helps develop existing connections to music and generate new ones?

## FRAMING THE PROBLEM

In thinking beyond user needs and Jobs To Be Done (see, for example, Ulwick 2016), we were inspired by the sociologist Lucien Karpik and his book *Valuing the Unique: The economics of singularities* (2010). In this work, Karpik argues that cultural products such as music, wine, novels and movies, are singularities – complex, multidimensional goods, the value of which can't be reduced to their specific features. It would be silly to claim one song has more value because it is longer, or because the singer hits higher notes. Or that a glass of red wine should be more expensive because it is a darker hue. Focusing on features in isolation misses the point.

Because value cannot be easily assigned to singularities, markets of singularities rely on complex mechanisms that enable actors to make decisions and choices and navigate uncertainty. Whereas in markets of commensurable goods, actors compare costs and benefits, in markets of singularities, they rely on what Karpik calls judgement devices and trust devices. Judgement devices "act as guideposts for individual and collective action" (Karpik 2010, 44) by providing cognitive support and opinion. Examples include reviews, charts or personal recommendations (Karpik 2010, 44–54). Trust devices help remove, dissipate or suspend uncertainty (Karpik 2010, p. 56) because they are often part of larger symbolic systems, such as social norms or formal authority. In the case of singularities such as movies or wine, we rely on the movie critic or wine connoisseur (and their training, education or expertise) to tell us what to expect, guide us in refining our tastes, teach us how to articulate the nuanced differences in our experience and ultimately, they help us make judgements about what is good and what isn't, what we like and what we don't. Because of the cultural complexity of singularities, we rely on these devices to serve as proxies of value.

Music is a type of singularity. It's a type of product that requires knowledge and tastes for us to be able to make a judgement and choice about what music to listen to. Historically, radio has played an important role for the segment of listeners who are not confident in their

ability to make choices about music by offering listening experiences curated by radio hosts who navigated the uncertain cultural field for audiences while also providing non-musical engagement, entertainment and information.

To continue differentiating itself as the leading audio streaming platform, Spotify needed to find ways to become a better 'judgement device' and a 'trust device' for songs, to use Karpik's terminology, and to be able to do this in an automated way. However, to do this, we needed to move away from thinking about user needs and to understand what a meaningful experience of music can be.

## Needs Versus Meaning

The true meaning of singularities only emerges to the user when they experience them themselves. Unlike user needs or goals, the value of singularities can't be fully anticipated in advance. For example, Spotify's previous research using the Jobs To Be Done framework, had identified that most listeners listen to music for fulfilling one or more Jobs (e.g. helping them focus, helping them change their mood, helping them create an ambience etc.). However, the Jobs To Be Done framework does not help in understanding how the value emerges for the listener as the listening experience progresses. For example, two users may derive entirely different meaning from the same good. One person could connect to a song because it soundtracked a breakup while someone else could love the same song because it gets people dancing. Jobs To Be Done framework and need-oriented approaches in general, miss this very important nuance.

Karpik explains that this uncertainty about what is valuable, which is a result of the incommensurability of cultural goods, is the defining characteristic of the market of singularities and requires an entirely different approach to value. In markets of commensurable goods, a consumer, Homo economicus, can make choices based on their needs and the expected costs and benefits. Their satisfaction is then derived in terms of efficiency. In markets of singularities, "Homo singularis must juggle the discovery, interpretation and evaluation of judgement devices; the discovery, interpretation and evaluation of singularities; sometimes the discovery, interpretation and evaluation of his own tastes; and a reasonable use of scarce resources." (Karpik 2010, 67)

Following Karpik's distinction between a need-oriented approach and a meaning-oriented approach has enabled us to come up with an entirely different model for thinking about the role that Spotify needs to play for its users and how this should be scaled throughout the organisation. We suggest that a meaning-oriented approach is more suitable for application to any services that offer cultural goods, such as music, video, film, fashion and luxury products, because it opens up opportunities to provide not only personalized experiences but also more relevant and meaningful experiences.

To illustrate the difference between the two approaches, consider a listener who may want to listen to the song *Run the World* by Beyoncé to improve their mood and feel motivated. Following a need-oriented approach, the job to be done is to enable them to search for the song, find it and play it as quickly as possible and without unnecessary friction to avoid frustration. Following a meaning-oriented approach might reveal that the listener experiences the song as meaningful because they identify with the archetype of a strong woman that this song represents. The recommender system could then songs by other artists who represent the same archetype, such as P!nk.

This distinction between a need-oriented approach and a meaning-oriented approach has implications for research design and analysis. While a need-oriented approach benefits from focussing on jobs to be done, the meaning-oriented approach benefits from the exploration of various judgement and trust devices that people use in order to navigate singularities. These can include recommendations from friends, popularity of artists but also personal aspirations. memories, travel experiences or social norms. A quote from Ken, 32, illustrates the complexity of music experience. Ken cannot easily identify his need or a goal. The meaning of the experience unfolds as his music is enjoyed by others, he gets complimented on it and helps him make new friends.

> 'Honestly it's whatever sounds good to me, that I can imagine myself at the beach listening to it I put it on, or anything that sounds good that I think other people will like… It creates a nice vibe, in the beginning I like it because it hypes you up to volleyball. A lot of people have said my playlist has been good, or they like my playlist and this is one woman in the group keeps on saying, I love your music because she use to be a DJ and then she is like oh! Can you share your playlist… I try to incorporate music that everybody likes so I have everything in there, and it's constantly updated… It's good for making friends, makes everyone happy.'

The proposed meaning-oriented approach suggests how we might help users like Ken make choices so they can have a more meaningful experience on Spotify.

**Table 1. Comparison of needs-oriented analysis and meaning-oriented approach**

| Needs | Meaning |
|---|---|
| Pre-existent - they drive choice | Emergent - can't be anticipated in advance |
| Binary - are either met or not | Multiple - experienced in multiple, unpredictable forms |
| Choice is rational and based on calculation of expected costs and benefits | Choice and actions are justified when meaning is present |
| Meeting needs does not affect identity | Meaningful experiences affect individual's identity |
| Systems are judged on efficiency | Systems are judged on the quality and relevance of the meaning |

*The table is inspired by Karpik*

Beyond User Needs – Hajdakova et al.

## METHODOLOGY

To tackle this challenge, we conducted ethnographic research with 12 participants in Boston in 2019, representing 2 types of audience segments:

1. **Lean in users** are knowledgeable about music, understand their own tastes, have the vocabulary to articulate their preferences and are confident in discovering new music. These users understand musical genres and remember artists and songs.

   An example of how a lean in user can express his musical preferences:

   *Listening for the beat, I'm listening for the lyrics, what the artist is actually saying – do they flow on the beat, does it sound good together. Yeah, pretty much a good beat and then like good lyrics can win me over if executed well. (Noah, 25)*

2. **Lean back users** are not confident in understanding established categories, such as genres, nor are they confident in their own tastes. They struggle to remember or articulate what they like and rely on others to help them discover new music.

   An example of how a lean back user expressed her attitude towards music:

   *I'm more a radio person. So when it comes on the radio I'll listen to it but I would say music is not something I'm like super obsessed … Like I enjoy music, I like it. But some people are always listening to music and always want to search for their own music, create their own playlist like they have a particular taste whereas I'm very much like fine with usually what's on the radio. (Jennifer, 30)*

Recognizing the importance of lean back listeners for further growth, we over-indexed on this segment.

Prior to 3-hour face-to-face in-home interviews, we engaged respondents through mobile diaries, asking them to report on at least 3 instances when a piece of music stood out to them and they experienced it as meaningful. We asked them to capture these settings and describe how they felt and why. In interviews we further probed into these and other instances to understand how people connected to music and how they experienced it as meaningful.

Key to the project's success and organizational impact was the engagement of stakeholders and various teams and working closely with other researchers. We conducted stakeholder interviews and invited members of the Spotify team to join the fieldwork and asked them to share brief reflections on each interview immediately afterwards. These videos were then shared throughout the organization to engage more people.

During fieldwork, at the end of each day, the team gathered together for a debrief. We put up posters with photos of participants and their homes and captured their most important musical connections. At the end of the download session, we recorded brief videos about each participant and their meaningful musical experiences that were shared via Slack with the wider team.

This process provided ongoing interest in the project in all its phases - from fieldwork and insights to organizational and product implications and further opportunities for engagement between Spotify and Stripe Partners.

## INSIGHTS

Altogether, we gathered more than 400 instances of significant connections to music and analyzed them using Airtable to generate insights.

### Two Cultural Engagement Models

We identified 2 cultural engagement models that listeners use to relate to music and experience it as meaningful: (1) musical engagement during which <u>music</u> is the focus of the experience; and (2) non-musical, during which the <u>listener</u> is the focus of the experience.

**Musical engagement** tends to utilize established cognitive tools such as universal vocabulary describing musical properties, classification into genres and historical periods. It draws upon expert opinion and values uniqueness and originality.

> Like Kanye West has the song Runaway that came in 2010 and all he does at the beginning of it is, he hits the E6 key, hits it again, hits it again and that goes on for like 15 seconds, he then goes down to E5. Like that one drop in octave, that was so cool for me when I learnt how to do that. (Rodrigo, 25)

**Non-musical engagement** uses more 'fuzzy' classifications often derived from personal experience, rather than universal categories, and tends to be attuned to popularity and common opinion and enjoys relatability of the artist or the song and the sense of being similar to the listener.

> At work, I like to listen to Beyoncé. That's another R&B. Her songs are—they're appropriate. I like listening to them. ... I like Beyoncé because you relate. One of my favorite songs from her is Run the World, how women run the world. (Yasmin, 32)

These two cultural engagement models are products of culture and shape the way we experience and enjoy music and express our passions for music. Both models are available to us and, as the research revealed, people sometimes switch between these models or enjoy music through a combination of both models. For example, we met participants who considered themselves experts in music capable of articulating their tastes in niche genres but there were songs they mostly enjoyed because they reminded them of someone else.

The two cultural engagement models provided more nuance to our understanding of the two audience segments and their ways of listening. We understood that it is not our goal to teach the less knowledgeable users about music in order to nudge them into the musical type of engagement. The lean back listeners were well aware that the musical engagement model was available to them and often had someone in their life who was a lean in listener, however, they did not necessarily want to enjoy music in the same way. For example, Maria,

28, grew up in a strict religious environment and was only allowed to listen to Christian music. Now, she doesn't want to feel any obligation to listen to music in any specific way.

> 'So I grew up in a very strict religious background and there we were not allowed to listen to secular music and could only be religious music and at that time they were like made you focus on the lyrics… I feel like they made you analyze everything single thing and I think that's why now I don't wanna analyze. Like I just want to listen and enjoy.' (Maria, 28)

Lean back listeners like Maria were capable of enjoying music just as much as lean in listeners, however, they often experienced difficulties and lack of confidence in their abilities to find the music they liked and discover new music, which is why they regarded radio as useful. This insight revealed the importance of understanding how Spotify might be a better 'judgement' and 'trust' device for those listeners who default to the non-musical model of engagement.

## Experiences Of Meaning

Inspired by Karpik, in our analysis of the instances of meaningful connections to music, we looked to understand what helped people make decisions and judgements about music and ultimately, how they derived value from listening. No matter what audience segment people belong to or what type of engagement with music they have, they all look for meaning and meaningful listening experiences. However, listeners, especially lean back users, don't necessarily know where and how to look for it. When they rely on automated recommender systems, they fear ending up in their own echo chambers or with irrelevant suggestions. While we cannot anticipate what they will experience as meaningful, we can give them cognitive and evaluative aids to enable them make choices and validate their efforts.

This approach revealed 9 distinct types of experiences of meaning that people have with music, each defined by specific types of tools and aids that we call *cues* and *proof points*. Cues provide a sort of navigation towards an experience of meaning and help locate meaning. Proof points help affirm the meaning, validate it and enhance it. An example of a useful cue for Yasmin, who identifies with Beyoncé might be knowing that her song *Spirit* celebrates Africa. A proof point then would be a video featuring African fashion.

*The Nine Experiences of Meaning*

The following are the 9 experiences of meaning. We illustrate the differences between the two cultural engagement models and the cues and proof points that help them navigate their experiences.

Cultivating knowledge | Differentiating tastes | Participating in | Exploring cultures | Identifying with | Remembering moments | Sharing with | Adapting to social context | Being moved

1. **Cultivating Knowledge.** Music is experienced by building a pool of knowledge and reflecting an existing pool of knowledge, including knowledge about how music is created, how it can be listened to, and how it varies across geographical areas and historical periods.

   **Musical engagement** derives meaning from enjoying and developing musical expertise. Listeners locate meaning through established categories, e.g. genres, expert opinion or their own musical training. Meaning is affirmed through proof of authenticity and uniqueness of music.

   *[On the decision to learn to play the piano] It's totally personal because I am not gonna do anything with the piano side of it, but it's kind of cool just to learn it. Again, for me being someone who loves music, I just like to play the song as I like to listen to. (Rodrigo, 25)*

   **Non-musical engagement** derives meaning from connecting music to another area of life or type of art, such as recognizing the aesthetic style of an album artwork. Listeners use general categories (e.g. the 2000s) and friend recommendations to navigate music and enjoy knowing a particular song represents something bigger or is popular.

2. **Differentiating Tastes.** Music is experienced by perceiving the scarcity and heterogeneity of music.

   **Musical engagement** derives meaning from feeling that one's taste is unique and desirable. They use expert validation as a cue and being asked for recommendations is a meaningful proof point.

   *Yeah, if there's anybody who could talk music just as much as me it's him. And so like how people go to me I'll go to him for artists that aren't well-known and stuff like that. So, it's always Greg Put You On and it's always these smaller artists that aren't as well-known. (Noah, 25)*

   **Non-Musical engagement** derives value from music when their taste is acceptable and shared with others. They use common opinion to make decisions about music and get more value out of the experience if they realize that their tastes are shared by others.

3. **Participating In.** Music is experienced as a unique encounter with the musician(s) and an embodied experience of music, shared with others.

**Musical engagement** derives meaning from being able to better understand music and one's tastes through the embodied experience. Scarcity and rarity of performances is an important cue.

**Non-musical engagement** finds meaning in the feeling of intimacy and proximity to the artist. Listeners enjoy the sense of the experience being shared with others.

*My husband introduced me to [a Haitian band] when we were dating and then I met the lead singer in person and I met the whole group when they were together. So, it's like I get different emotions because I met them. (Yasmin, 32)*

4. **Exploring Cultures.** Music is experienced as a novelty and a comprehensive representation of the particular scene (e.g. popular music, specific genre, specific "terroir" etc.)

   **Musical engagement** derives meaning from developing musical expertise in a new culture (or sub-culture). Listeners orient themselves through established categories but through a different context. They enjoy the novelty and discovering something unknown or hidden.

   **Non-musical engagement** derives meaning from being able to identify music with travel experiences and other personal interests. Important cues include their past travels and recognizable cultural signifiers (food, dance, etc.). Value of music is reinforced when this can be relevant to others, such as family or friends.

   *I'm the only one who really knows how to dance Copa music. I encourage my little sisters they know, my little brother not at all. He knows nothing about the culture. … And then with having the kid, I'm like now I really like to learn to speak creole, learn to love Haitian food and try to know how to at least dance the little Copa two step, you know that way you're not completely lost if you were to go to Haiti, you know people will be like, "Oh, he's a Haitian kid" (Tamila, 29)*

5. **Identifying With.** Music is experienced as a language, expression, mental shortcut and visualisations that anchor and elucidate personal narratives and aspirations.

   **Musical engagement** derives meaning from sensing that music aligns to both tastes and personal aspirations. Listeners love artists who are central to a wider culture or a sub-genre and value unique stories.

   **Non-musical engagement** derives meaning from being able to aspire to the lifestyle the music or artist represents. They recognize desired archetypes and enjoy being able to identify with the wider lifestyle of the artist, including their fashion and values.

*When I was listening to [the new Beyoncé song], I also looked for the video, and it made that connection of the relationship that she has with her oldest daughter. So, it brought me to the like, oh, I kind of want to have that relationship with my kids… (Yasmin, 32)*

6. **Remembering Moments.** Music is experienced as a bridge to anchor memories, provide shared reference points, stories and exclusive tacit language.

   **Musical engagement** gives memories a musical significance. For example, listeners enjoy when their lives correspond to important milestones in the history of music or in the careers of their favourite artists.

   **Non-musical engagement** likes using music as a shortcut to one's own personal memories. Listeners locate the meaning of music in their own memories of places, times and people and love sharing memories with others.

   *It helps me remember the people that were in my life, so, Summer camp people were awesome. So, it transports me back to the feeling of safety and being part of something bigger than yourself… (Lilly, 33)*

7. **Sharing With.** Music is experienced as part of a shared experience and serves as a shared reference point.

   **Musical engagement** enjoys deepening their musical experience with people with similar tastes. They use shared vocabulary and enjoy a proof of shared niche tastes.

   **Non-musical engagement** derives meaning from music enabling them to deepen relationships through shared musical references. They use shared activities, such as cooking, or family trips as important cues and love to see a proof of others' tastes and moods.

   *Going to the concert was like nice because it was a bonding experience and we were there for kind of a longer amount of time, because I bought this Airbnb. And then we were just sort of together and it was like a special shared experience, and then just like music is tied into that.' (Sylvia, 18)*

8. **Adapting to Social Context.** Music is experienced by creating a vibe while minimizing social friction.

   **Musical engagement** derives meaning from the alignment between one's musical tastes and the social context. Listeners enjoy seeing a proof that other people have the same tastes and express enjoyment in a shared way.

   **Non-musical engagement** focuses on music being appropriate for the occasion. They locate meaning in social norms and enjoy seeing positive social responses.

   *'Sometimes we'll listen to music like if we're like entertaining or outside or the kids want to play in the yard or something like that, we're able to drink a beer. We'll listen like out on an outside*

*speaker… Things like '80s music, Billy Joel, like you can listen to that around anyone, kids, neighbors without worry about you know offending them… I think outside is more like Jack Johnson and not Drake radio because the neighbors can actually like hear the music so I put a genre here that's more appropriate.' (Max, 33)*

9.  **Being Moved.** Music is experienced by changing the physical or emotional state.

    **Musical engagement** derives meaning form the alignment between one's musical taste and the activity their doing or the state of the mind they're in. They can explain why they feel moved using musical terms, e.g. specific tempo that gets them motivated, and they can select specific music they know will work for them.

    *'Music is integral to working out in the gym. I couldn't workout without music. It's motivating. It's the equivalent of having like a couple of cups of coffee. I work harder when I'm listening to music, it's motivational in a way… I like loud hip-hop music, a quick beat, aggressive lyrics… Yeah. It's motivating and it pushes me more than I would push myself if I wasn't listening to music...It's motivating and if I didn't have it, I'm not in charge of the gym.' (Trevor, 34)*

    **Non-musical engagement** derives meaning from music that is supportive of an activity and navigates music based on the type of activity or mood, e.g. searching for 'a running playlist'

## DESIGN / PRODUCT IMPLICATIONS

This framework provided a new language and a new foundation for the wider team to think beyond user needs. It also helped understand the role and importance of various types of tools and media that have been an important part of music industry but have not necessarily been successfully embraced by streaming services. For example, the framework explains why listeners who default to the non-musical type of engagement often prefer the radio or why they like YouTube. These services provide them with the cues and proof points they are looking for, such as popularity charts, imagery and comments from other listeners. The insights have enabled Spotify to critically reflect on the fact that the app was more suitable to listeners who default to the musical type of engagement and have given a useful direction for improvement.

The learnings and the framework had direct implications for product and design. The most straightforward application of the framework was for the user interface to use cues and proof points to locate and amplify meaning for the audiences who default to the non-musical type of engagement. For example, possible product interventions in the future could include personalized cover images and annotations.
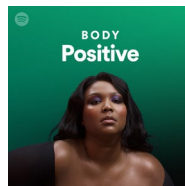
For example, Yasmin, 32, who liked to listen to artists she could identify with, (e.g. with Beyoncé as a mother, or with P!nk as a strong woman who's 'her own boss') might enjoy being introduced to Lizzo. Useful cues can include visuals celebrating Lizzo's body positivity and a proof point might be a playlist that offers a gateway to discovering more artists who celebrate body positivity.

The research also helped identify new ways for recommender systems to forge meaning. Meaning is not fixed, it fluctuates depending on context and also changes with time. While meaning cannot be anticipated, we can still aim to optimize it by utilizing cues and proof points. Cues and proof points can help intensify existing connections and add more meaning to them by helping listeners connect to multiple experiences of meaning. For example, Yasmin might enjoy listening to a playlist that Beyoncés's shared with her daughter.

## ORGANISATIONAL IMPACT

The success of this project relied on the ability to scale the framework throughout the organization. This was a crucial part of the project from start to finish. Stripe Partners conducted multiple stakeholder interviews in order to understand the organizational context of the challenges but also to engage stakeholders in the project. We also used Slack to share interesting links and provide connection between all the researchers involved across all research streams.

At the end of the project, Spotify hosted a share out session with teams joining in person and remotely. We presented insights based on ethnographic evidence, stories from the field and videos from the diaries and used them as a basis for the final framework. The final session also included activation workshops to get participants to think of immediate takeaways and to apply the framework. The session included 'office hours' when anyone from the organization could learn more and discuss the project with the researchers. We created visually engaging outputs in form of posters and also did a video recording of the presentation for Spotify's archive.

Within Spotify the learnings from the research has had a broad impact across multiple business units and has been internalized deeply into Spotify's culture.

From an agency point of view, this project was relatively unique in the sense that we were engaging with the client's stakeholders, teams and other researchers and experts from start to finish ensuring the insights and the framework were relevant and easy to understand. This has allowed us to understand the organization better and have conversations with people working on diverse challenges to help them think through potential implications for their areas and goals. Ultimately, this enabled our work to travel and develop new client relationships.

## REFERENCES CITED

Karpik, Lucien. 2010. *Valuing the Unique: The Economics of Singularities*. Princeton and Oxford: Princeton University Press.

Ulwick, Anthony W. 2016. *Jobs to be Done: Theory to Practice*.

CATALYST

# Leveling Up Your Research and Research Operations

## Strategies for Scale

BRIGETTE METZLER, *Services Australia, University of Queensland.*

*With user research becoming more common within organisations, there is an emerging issue of meeting demand whilst also developing the craft of research. A new profession is emerging in response – research operations. This paper will describe the current state of publicly available frameworks for research operations. These tend to deal with one aspect of scale – the people who are doing the research, not how they do the research, when, or what we do with the research. Two frameworks will be combined to create a matrix that provides the tools to identify an investment strategy for research within the context of an organisation and their strategic goals. This matrix provides a significant contribution to the field by making it possible to be strategic and proactive about developing research practices in the  context of individual organisations, how and why they do research, and to better manage the tension between scale and craft.*

*Keywords: ResearchOps, UX, Strategy*

## INTRODUCTION

Qualitative research as an embedded practice in industry and in government has been emergent since at least the 1960's[i], and has grown to the point of being commonplace in the world today. Modern design is largely attributed to the collision of the arts and crafts movement with the machine age.[ii] Alongside this emerging popularity, the practice of modern design has matured, and our understanding of art and design as an embodied experience, one worthy of replicating in applied ways, has also matured. Seen in this light, it comes as little surprise therefore, that in recent years, the attention economy[iii] has raised the profile of (and pressure on) human researchers and research outcomes even further, as companies attempt to squeeze out every last minute in a person's day that can be spent on their device, in a platform, and/or watching ads. It is a bleak picture to paint, but the opposite story is there too – through human centredness/system centredness, researchers have a role in enabling industry and government to have a meaningful impact on people's lives. Our current state with regard to the health of the planet and the people therein may cause many to want to engage in qualitative research as a part of their design work in order to have the best possible chance of effecting meaningful, ethical and human centred change. That means it may seem that everyone, from the smallest to the largest company, are employing researchers to do more and more research in less and less time.

The possibility of effecting that change is incredibly exciting, intoxicating even. In many ways, the chance to do some real, lasting change work has never been more present. The profession of applied qualitative research (commonly known as user research, as it will be

referred to throughout the rest of this paper, but also including design research, UX, CX – largely a broad umbrella of qualitative, human led, conceptually post-modern research) has developed significantly in the past few decades. Alongside traditional ethnographic and anthropological or human factors research methodologies, or other disciplines such as psychology, researchers can be trained in systems thinking, and in the practice of co-design.[iv] It is possible to see that the field is becoming more established.

The reality of having teams of any size – whether 1, 100, or 1000[v], is that the demand for research far outstrips anyone's ability to meet that demand. While it might be traditional for a research team of one, or 5 even, to individually be running their own processes and procedures, their own contracts, panels and ways of working when it comes to research data management and sharing, it can become a huge time impost, leading to duplication, unintended replication and burnout.

The ResearchOps Community is a volunteer run online community of over 6000 (to date) individuals from 62 countries working in the field of user research and research operations, coming together with a common goal of giving shape to, and validating research operations as a profession. The common challenge facing the members of the community is doing research at scale.

As work has progressed on understanding what research operations is, so too has our understanding of how to manage the tension of delivering research at scale, whilst also maintaining rigour in research. This paper will describe the current state of frameworks (that are publicly available) for research operations, which, to date, tend to have been *maps* for research and research operations. Following this, the paper will bring together two frameworks that, once combined, allow one to see the *terrain* of research in *individual* contexts. This blending of frameworks, known as the Pace Layers Matrix is the result of observation and experience from 3 global research projects undertaken by the ResearchOps Community[vi] (one on what research operations is, one on building a research skills framework, and the most recent on research repositories) and the author's own work in understanding the research outcomes from those projects. Having the tools to identify the terrain of one's own research practice in context provides a significant contribution to the field by making it possible to be more strategic and proactive about developing research practices in context and better manage the tension between scale and craft.

## The Problem with Scale

By now, the industry is well and truly aware that there is a problem with managing the demand for user research. The issue is a seemingly simple one – qualitative research takes time, lots of it, and this does not scale well. Demand for research grows, and the expectation that good research can be done in months moves to weeks, and sometimes even days. At what point fast research becomes poor quality research is what is at issue. At what point the profession suffers from poor quality outcomes from overstretched, under-resourced or untrained researchers is an ever-present burden when the topic of scale emerges.[vii]

There are myriad ways to deal with demand – adding more and more researchers, creating hub and spoke models to have core researchers at the hub and 'people who do research' (PWDRs, a phrase coined by Kate Towsey[viii]) operating in small teams, or having a core group of researchers embedded individually across the organisation but reporting to a central research leader. But each of these models only deals with one aspect of scale – the people

who are doing the research, not *how* they do the research, *when*, or *what* we *do* with the research.

Research operations has emerged from this gap – a field dedicated to:

> "the people, mechanisms, and strategies that set user research in motion. It provides the roles, tools and processes needed to support researchers in delivering and scaling the impact of the craft across an organisation." [ix]

Within the field of research operations (also known as ReOps, or ResearchOps), there are a group of sub-fields, all addressing slightly different issues to do with how we create the right environment for research to happen. They include (non-exhaustively):

- Making better use of existing research through the creation of a research library or repository, though these often fail to achieve the results that are hoped for.
- Research operations playbooks or 'centres of excellence' are one of the first ways one sees operations leaders attempting to address the 'how' research happens at scale. This also attempts to address the additional issue of the tension caused when trying to do a lot of research in a short time – creating efficiency, and also enabling others to do research, even if not fully trained in doing research.
- Systematising and streamlining recruitment
- Centralising budgets and managing tools, platforms and contracts centrally within large organisations

All of these responses to scale create ripples that are felt across the organisation and the broader user research industry. Indeed, the concept of the democratization of research is a hot topic, occupying whole streams at research conferences (see for example: Advancing Research 2020[x]) and the topic of debate in blog posts and papers.[xi] Interestingly, the democratization of research has long been a topic of debate in academic circles also, but it is framed instead as a feminist act, or an act of 'research justice',[xii] enabling research to be decolonised. This is not two separate disciplines using the same terms for different ends. Rather, it is a different lens on the same issue – the practice of extending research spaces to people outside the role of research. From *researched* to *researcher*, from *consumer* of research to *doer* of research. Kara states that "The term 'democratizing research' covers a range of emancipatory approaches to research such as activist research, feminist research, decolonizing methodologies, community-based research and participatory research".[xiii] User research, being embedded within design and design principles such as co-design, co-production and others[xiv] is a practice of research in, of, and sometimes with, community, and is often participatory – the tension therefore, is the same.

## Complex Systems and Frameworks – Tasks vs Strategy

To understand the current state of 'research at scale', it is worth acknowledging the myriad frameworks that have arisen in the wake of the emergence of research operations. Given that the ResearchOps community (that really is a catalyst currently for the emergence of the profession and the development of frameworks for understanding what ResearchOps

is) started with the 'what' of research operations, early frameworks, such as the ResearchOps Community's 'What is ResearchOps'[xv] and the Nielsen Norman Group's 'ResearchOps 101'[xvi] have focused on dividing up the *tasks* or *things* that need to be done in order for research to happen effectively. It is possible to see however, that tasks are only part of the way research happens. *How* it happens, with *whom*, and most importantly for this paper, *in what contexts*, are all essential in effectively delivering research at scale.

## Towards a Strategy for Scale: PESTLE Models

Towards the end of 2018, a group on the board of the ResearchOps Community (Emma Boulton, Holly Cole, Tomomi Sasaki and myself) realised that the taxonomy, or the conceptual framework we'd applied to understand the data from the 'What is ResearchOps' project could be used to understand the relationship between research and operations. Emma Boulton took this forward with the 8 Pillars framework.[xvii] This model (Figure 1 below) can be seen as a typical PESTLE strategy model. The PESTLE model arose from the work of Professor Francis Aguilar following his book, *Scanning the Business Environment* in 1967.[xviii] It is a framework for understanding the political, economic, socio-cultural, technological, legal and environmental factors that are involved in managing business, with the idea being that if one is aware of the forces impacting the business, then it is possible to create a strategy for optimizing opportunities and mitigating risk.

| Research Ops | | | | | | | |
|---|---|---|---|---|---|---|---|
| People silos Education Value of research Buy in Internally focused

Push back

Stakeholders Executives Colleagues | Research as a team sport Cadence Sharing insights

Prioritisation Integrating insights Processes Methods Protocols | Community of practice Professional development Staffing

Leadership

Mature career paths

Org Design | Space Time Resources

Market forces Org Maturity

Budget R.O.I.

Business constraints | Scheduling Logistics Paperwork Timesheets Incentives Panel management Participant coordination | Document templates

Research library Data silos Data gardening Knowledge management | Risk assessments GDPR Legal Infosec

Consent Ethics | Technology Procurement Software Hardware

Labs Systems Networks |
| Environment | Scope | People | Organisational context | Recruitment and admin | Data and Knowledge Management | Governance | Tools and infrastructure |

@teamreops / @brigettemetzler

Figure 1: 8 Pillars of User Research

In the same way, the 8 Pillars, with the focus on *environment* (in the PESTLE model: environment), *scope* (political), *people* (socio-cultural), *organisational context* (economic and political), *recruitment and admin*, *data and knowledge management*, *governance* (legal), and *tools and infrastructure* (technological) can be used as a way to generate an understanding of the factors and forces at play when research happens.

Briefly stated, the 8 Pillars, as they pertain to research, are:

- Environment: Why does research happen? Who engages with what I do?
- Scope: The nuts and bolts. Methods, processes. How and when does research happen?
- People: Research maybe done by designers or product managers. Can we create a community of practice to support and mature the craft? What does a career path look like?
- Organisational context: What is the maturity level of the organisation I work in? What are the external constraints that affect me? Things such as budget, resources, time, space.
- Recruitment and admin: How do I manage all the logistics and admin for research projects and participants?
- Data and knowledge management: Often valuable insights are lost over time as teams grow and change. How do we ensure the same studies aren't repeated? What happens to the research findings, data and insights?
- Governance: As a researcher what are the legal and ethical considerations that affect my work?
- Tools and infrastructure: What tools and infrastructure do I need to help me with my research?

The 8 Pillars is an effective strategy framework and provides a high-level view of the concepts and things that need to be in place to make research happen. However, over time, as the community grew, it became obvious that different *methods* employed to do user research were also important to understand as the concepts and things that need to be in place to make research happen are dependent on the method employed. Across the industry, researchers tend to use mixed methods dependent on context – on time, resources, capability and of course, the research question and outcomes required. The 8 Pillars provides a *map* – but to take a step further, what's required, is knowing the *terrain*. The profession of research operations has moved from a list of what is present in the map (the 'what' of research operations), to a 2D map of research (the 8 Pillars here can be viewed as a way of seeing the pathways present in getting us 'down the road'). But what good is a map if it cannot tell us whether to pack a kayak or snowshoes? How big is the mountain? How steep is the road?

## Complex Systems and Pace Layers

To that end, Brand's Pace Layers framework can be used to see the terrain. In 1999, Stewart Brand wrote *The Clock of the Long Now: Time and Responsibility*. In it, he suggested Pace Layers could be used to understand complex systems.

> "Fast learns, slow remembers. Fast proposes, slow disposes. Fast is discontinuous, slow is continuous. Fast and small instructs slow and big by accrued innovation and by occasional revolution. Slow and big controls small and fast by constraint and constancy. Fast gets all our attention, slow has all the power." [xix]

This is deeply analogous to research, because Pace Layers are all about time, speed and depth. An important point about the above quote is that 'fast learns, slow remembers', also, noting that 'fast gets all our attention and slow has all the power'.

In many ways, organisations with sufficient research to have need of an operations function can also be understood as complex systems. They tend to have researchers who use varying methodologies, have people who do research (PWDRs), and people who need research as well as full-time researchers. All these people influence how research gets done, why and when.
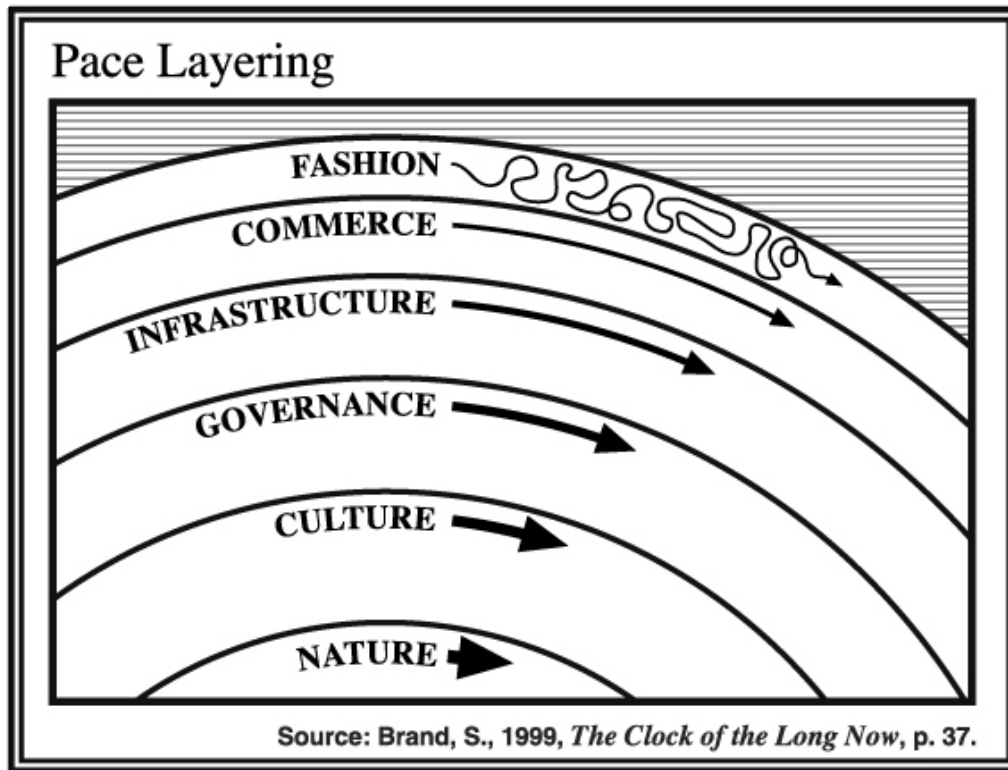


Figure 2: The Pace Layers Framework, © Brand, S, used with permission.

The Pace Layers shown above represent depth and pace. The deeper layers move or change more slowly, but conversely, also serve as a foundation. The higher the layer, the faster it moves and changes. Using Pace Layers to understand different research methods can help with strategy development as it takes a holistic approach to the ecosystem surrounding research. In this way, it is possible to move from a reactive position towards a proactive, strategic one.

## The Tension We All Feel: Constructive Turbulence

To Brand, the relationships between layers are key to the health of the system. Paul Saffo, a collaborator with Brand on the Clock of the Long Now project, goes further with this idea, stating that conflicts caused by layers moving at different speeds keep things

balanced and resilient. Saffo called this "constructive turbulence".[xx] Managing this constructive turbulence is the key to understanding inertia in the system, the things that constrain research teams, and the opportunities to scale. Turbulence in the system that is off balance can be seen clearly in the tension noted previously between the speed demanded by business and the time taken to do contextual, generative research. A symptom of imbalance, is researchers needing to spend so much time on faster evaluative types of research that they cannot gain the time or by-in for generative research favoured by ethnographic research methodologies for example. Or researchers spending so long doing the deeper layers of research that they are unable to respond quickly or lack the skills or infrastructure to do evaluative research when it is required.
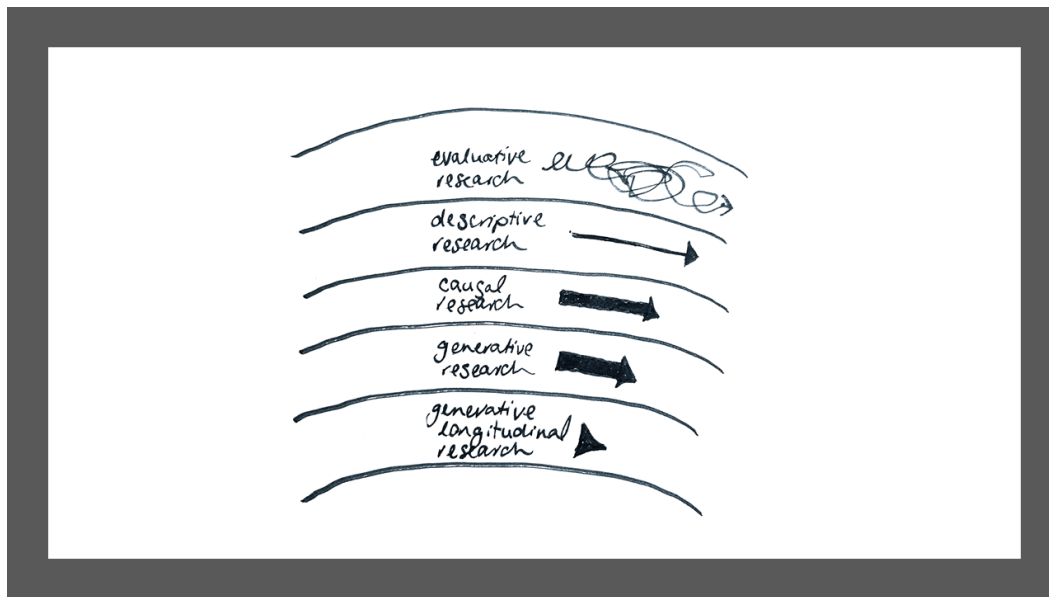


Figure 3: If research methods were Pace Layers. Pace Layers framework used with permission.

A consistent complaint of researchers who use slower and more in-depth research methods (those closer to the practice of ethnography and anthropology- closer to 'people' and further away from researching people in the context of 'things'), is that they struggle with constant pressure to reduce the cadence of their research. They struggle with pressure to deliver according to the cadence of business, rather than deliver within the traditional research methodology (lots of observation and research at the start, during which almost nothing is 'delivered'). If we use Pace Layers to understand the nature of generative research and its place within the system, this will help us to reframe the value of the slower, deeper layers, and also to see the friction between those layers.

On the other hand, the research that gets done at the top is the one that makes all the noise. There is high demand and constant pressure with regard to time and the findability of the insights generated in this layer. Researchers tend to be working within agile sprint cycles with product development teams. Researchers doing mostly evaluative research, struggle with the denigration of the value of their research. Sometimes, they have research leads or executives wanting to be more strategic or to get more from the research than is possible.

The noise and speed can make it hard to drill down through the layers and get support for the slower types of research methods that can contribute to those strategic needs better than the evaluative research can.

## Pace Layers and Research Operations

When research methods are viewed using Pace Layers, it is possible to see that the operational work that needs to be done to help researchers do their best work is also different. All research methods require every aspect of the 8 Pillars to be in place in order for research to occur, but the focus is different depending on the dominant method used.
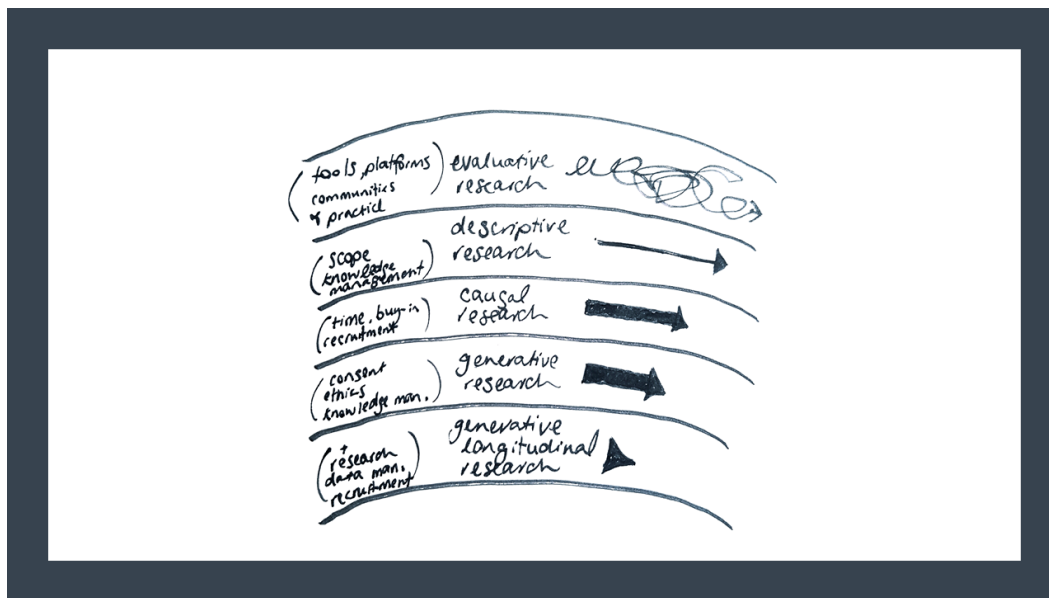


Figure 4: Research methodologies with their Ops foci — aligned with Pace Layers. Pace Layers framework used with permission.

In Figure 4 above, it is possible to see that the research methods used to do evaluative research will have an operations focus of tools and platforms for research such as moderated and unmoderated usability testing. Evaluative research tends to be more frequently done by researchers or PWDRs embedded in product teams, and so communities of practice become essential in ensuring research practices are consistent and that researchers can develop their research practice as a group. Descriptive research focuses on scope and knowledge management as it tends to be desktop research. Finding research that has already been done to evaluate is crucial. Having access to that research is crucial, as is receiving help with refining the scope of the research. Causal research has a focus on time, by-in and recruitment. Causal research requires a particularly rigorous approach which means that it can take a long time. Getting buy-in for these methods, such as A/B testing, is important, because it is the first of the layers in the model that really cannot be quickly carried out without undermining the rigor of the research. A careful approach to recruitment is essential

in order to ensure the validity of the research. Generative research has a focus on consent, ethics and knowledge management. Generative longitudinal is the same but also has a focus on research data management and recruitment. Ethnographic and anthropological research tend to use observational methods and require what Clifford Geertz famously described as 'deep hanging out'.[xxi] This generates a tremendous need for careful research data management (as this deeper, more contextual work tends to automatically accumulate a lot of personally identifiable data that grows in complexity and risk over time) including a long view on the ethics of the research and on ensuring participant consent is genuine, informed and that the participant retains the agency to work with the researcher on the way their data is used and managed *over time*. Research data management is important always, regardless of the research method, as is compliance with GDPR, however, it is simply that the complexity of managing this over time only increases as research data for an individual participant builds and layers over time. Triangulation of the data can render previously de-identified research data in practice, identifiable. Effective use of this growing parcel of unit level data in turn generates a large reliance on knowledge management practices.

## BRINGING THE MODELS TOGETHER: THE PACE LAYERS MATRIX

Understanding the layers and the 8 pillars as a matrix provides a tool for diagnosing strengths and weaknesses of a research practice and operations practice within an organisation, and therefore provides a path to both scale research within an organisation and also deepening research and operations capacity between the layers. To describe this effectively, two case studies are presented below. The first is an organisation with a strong research practice with researchers focusing on ethnographic (generative) research methods. The second is an organisation with a focus on utilising user research to best understand the use and effectiveness of the platform the organisation sells. Here there are a group of researchers embedded in product teams tending to work with developers through design and delivery to best evaluate the product development. The focus here is on evaluative research. The two matrices are shown side by side at first so that it is possible to see how the layers impact on the operations needed and how these vary in different contexts.
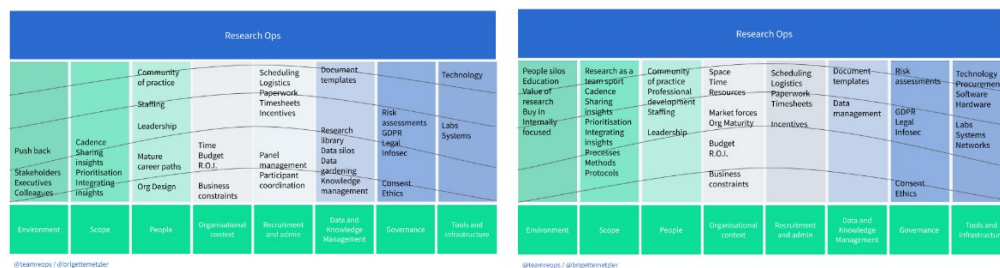


Figure 5: The Pace Layers Matrices in two organisations (explored in more detail below). The organisation on the left is case study 1, with a focus on ethnographic research practice. The organisation on the right is case study 2, with a focus on evaluative research. Pace Layers used with permission.

**Case Study 1**
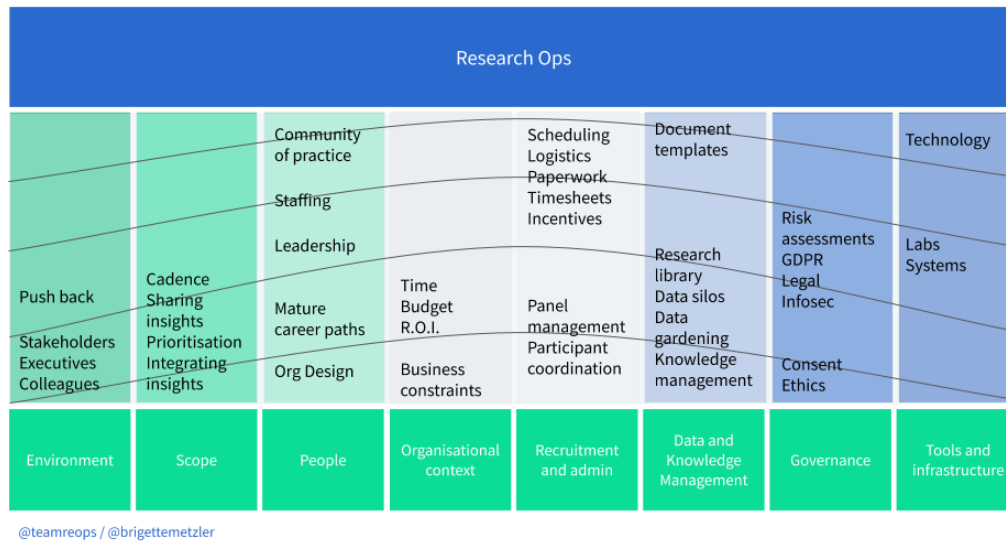


Figure 6: The Pace Layers Matrix in an organisation with a strong ethnographic research practice. Pace Layers used with permission.

The first case study is a research team of around 50 with a strong ethnographic research practice. The team tends to focus mainly on doing generative longitudinal, generative, causal and descriptive research. The focus in the *environment* pillar is on communicating with stakeholders, in the *scope* pillar on integrating previous insights, in the *people* pillar on developing the career paths of researchers, in the *organisational context* pillar on working within business constraints, in the *recruitment and admin* pillar on undertaking effective panel management and participant experience, then in *data and knowledge management* the focus is on research data management, in *governance* it is on consent, ethics and to some extent, in *tools and infrastructure*, having spaces and systems in place to allow researchers to be in the field. Their challenges will be about continuing to show the value of what they're doing, though they won't need to evangelise research as such. Instead, the push back will be the time it takes, the cost, it will be managing the melody of long and slow with the needs of business. They will do that through a rigorously managed panel, good participant experience, and by building their base of research to a level that others can dip into it as needed. Their research is very manual, so there is not a lot of focus on tools.

*How Might They Scale?*

The team's strength is in the depth and power of the deeper layer of research that they are creating for the organisation. If they can focus on getting the most out of that layer of research for the organisation through a research library, they can support the organisation to be able to move quickly (but with deep certainty regarding their evidence base) by giving the PWDRs a rich research asset to refer to. This will support less experienced researchers to feel more confident in their findings. They can be the foundation on which everything

grows. If the research team feels the need to utilise faster, less contextual research methods, moving straight to the top layer (evaluative research methods) isn't going to be effective, because they don't have the tools and technology in place to do so. Instead, they can direct their efforts towards moving gradually up the layers from the bottom (moving more heavily into descriptive research), or they could bring PWDRs into their research team who may have the tools and technology in place to do evaluative research, and focus their efforts on developing their research practice until the whole ecosystem has strengths across the layers.

## Case Study 2
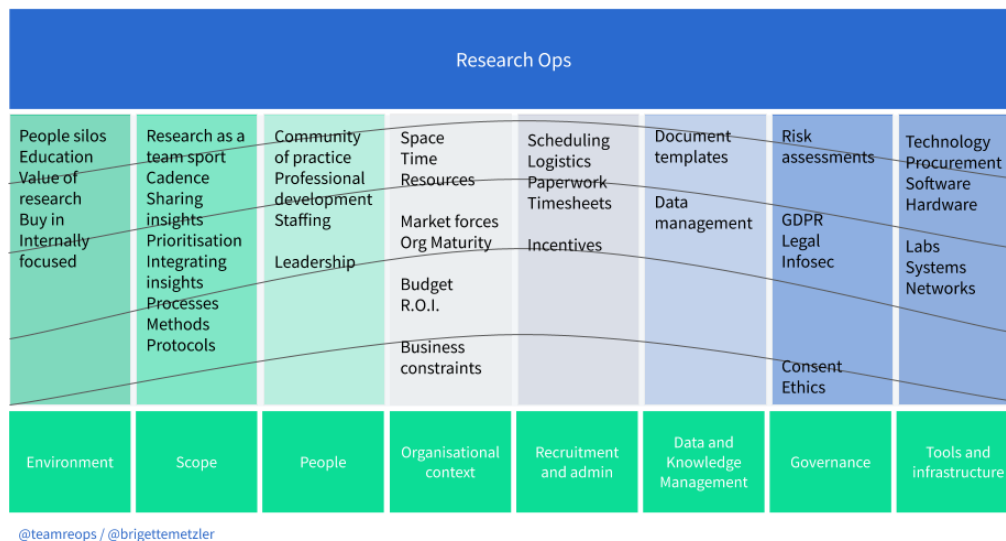


Figure 7: Case Study 2: Pace Layer Matrix for an Organisation with a Focus on Evaluative Research. Pace Layers framework used with permission.

The second is an organisation with a focus on evaluative research. This organisation has PWDRs and researchers dispersed throughout the organisation. Their focus in the *environment* pillar is on being careful to break down silos caused by being embedded across different teams and on gaining support to do deeper, more strategic research. In the *scope* pillar, Prioritising the research throughout sprint cycles is important, and in order to get a lot done in short time frames, the researchers will likely be good at treating research as a 'team sport'. This helps them continue to improve the buy in for more research. The design and development teams will expect insights to be delivered quickly, and their involvement in the research may generate tension around the democratisation of research. They might have descriptive and some generative research, but it will all be in support of the top, noisy layer.

Unmoderated usability testing and other methods undertaken in evaluative research requires a heavy focus on the *tools and infrastructure*. In the *organisational context* pillar, *recruitment and admin*, *data and knowledge management* and *governance* pillars, there will be a need for resources in the form of tools, templates, and guides for the PWDRs. In the *people* pillar, a community of practice could help them develop their research practice in the organisation. Given the research methods at play here, their consent is lightweight, most of the research is

de-identified right from the start. They will not be thinking about a library, and if they do, it will be likely held in whatever system the developers use to track their work.

*How might they scale?*

This organisation will struggle to develop their research practice without employing more experienced researchers. If their next hire is a senior researcher, their role could be to mentor and train the existing PWDRs. An evaluation of the skills of the current PWDRs is likely to uncover some people with skills in statistics due to the more quantitative nature of evaluative research methods. This could assist the organisation to undertake descriptive and causal research without significant change in the structure of their research teams, investment in technology or a significant increase in operational responsibility.

The barrier for this organisation in developing their research practice to include all layers is that they will need to be careful to communicate about the slow, deep layers and how they fit in the research lifecycle or the turbulence between the layers will become too great – the pressure on the lower layers to move at the same pace as the top layers will undermine any efforts they make to increase the depth. If they scale too quickly, panel management and data management (in terms of ethics and consent) will become a problem. If this organisation has an operations function, the Ops team can focus their attention on working with the researchers on enhancing research data management and panel management. If the team begins to notice the way they conceptualise their research has changed to thinking about their research as an *asset*, rather than as *evidence* (as people tend to do with evaluative research), then this will be a good indicator that their research practice has matured across the layers, and they are well on the way to a research practice that includes each layer.

## Pace Layers Matrix: Understanding the Terrain

The case studies shown above highlight that each organisation has a different context, but by bringing the 8 Pillars and Pace Layers Frameworks together, it is possible to chart the terrain within a single organisation – to identify individual strengths and weaknesses within the context of the organisation. Identifying the bumps in the road, the rivers and streams, mountains and valleys within an organisation is a difficult piece of work, best done together with all across the organisation who might be involved or have a stake in research. These people may be people who use research, people who do research, people who read research to aid their own research, operations people and people with a strategic responsibility in the organisation. Interviews using the matrix as a reference alongside a research lifecycle view can help uncover what each person perceives as the organisation's strengths and weaknesses, what they felt the organisation needed from research, where the demand for different research methods comes from, and what capacity they have to make the turbulence between the layers more constructive.

Some responses that people within the ResearchOps Community have used include co-creation days where the participants worked with product owners in real time on their products. Others have implemented a schedule for a stakeholder to join them on an observational session once a fortnight. A common response of course, is to implement an operations function to work on getting the 8 Pillars in place across research that is happening in an organisation. There are no simple answers to managing the turbulence

across the layers – it is an ongoing process of adjusting the threads, the warp and weft of the complex system, but it provides the mechanism to not only see the map, but the terrain of the system one works within to see the turbulence, and adjust the tension as needed.


## NOTES

i. Jones, & Thornley. (1963). Conference on design methods: papers presented at the Conference. Pergamon Press.

ii. Raizman, David. (2003). History of modern design: graphics and products since the industrial revolution. Laurence King Publishing, London, pp. 66.

iii. The concept of the attention economy stems from the work of Herbert Simon who stated in 1971 "A wealth of information creates a poverty of attention and a need to allocate that attention efficiently.", Simon, H. A. (1971) "Designing Organizations for an Information-Rich World" in: Martin Greenberger, Computers, Communication, and the Public Interest, Baltimore. MD: The Johns Hopkins Press. pp. 40–41.

iv. McKercher, K. A. (2020). Beyond Sticky Notes. Sydney: Beyond Sticky Notes.

v. Sekai, Farai, (2020, September). Twitter post, retrieved from: https://twitter.com/SekaiFarai/status/1308813419173347331?s=20

vi. The findings, outputs and talks on each of these projects can be found on the community website, https://researchops.community ).

vii. See here for example, Sirjani, B. (2020, July 22). *Democratization is our Job*. Retrieved from Yet Another Studio: https://yetanother.studio/blog/democratization-is-our-job, and Saha Mitra, S. (2020, April 22). Undemocratising User Research. *UX Collective*. Retrieved October 3, 2020, from UX Collective: https://uxdesign.cc/undemocratising-user-research-6897e6f4cf72

viii. Towsey, K. (2019, July). An interview with Kate Towsey, ResearchOps Manager at Atlassian. (N. Francis, Interviewer) Marvelapp. Retrieved from https://marvelapp.com/blog/kate-towsey-researchops-atlassian/

ix. ResearchOps Community. (2019, February). ResearchOps Community: About. Retrieved from ResearchOps Community: https://researchops.community/about

x. Rosenfeld Media. (2020, October 10). *Advancing Research Program*. Retrieved from Advancing Research 2020: https://rosenfeldmedia.com/advancing-research-2020/program/

xi. See for example Reichelt, L. (2019, November 11). Five dysfunctions of democratised' research. Part 1 - Speed trumps validity. Medium. Retrieved October 3, 2020, from https://medium.com/@leisa/five-dysfunctions-of-democratised-research-part-1-speed-trumps-validity-1606fa63a3e8, de la Nuez, A. (2019, August 12). Democratization of UX Insights: What does this really mean? *UX Matters*. Retrieved from https://www.uxmatters.com/mt/archives/2019/08/democratization-of-ux-insights-what-does-this-really-mean

xii. (2018). Democratizing Research in Practice. In R. Iphofen, & M. Tolich (Eds.), The SAGE Handbook of Qualitative Research Ethics (pp. 103-113). ProQuest eBook Central: SAGE Publications.

xiii. (2018). Democratizing Research in Practice. In R. Iphofen, & M. Tolich (Eds.), The SAGE Handbook of Qualitative Research Ethics (pp. 103-113). ProQuest eBook Central: SAGE Publications, p.103.

xiv. McKercher, K. A. (2020, October 3). What is co-design? Retrieved from Beyond Sticky Notes: https://www.beyondstickynotes.com/what-is-codesign

xv. ResearchOps Community. (2019, February). ResearchOps Community: About. Retrieved from ResearchOps Community: https://researchops.community/about

xvi. Kaplan, K. (2020, August 16). ResearchOps 101. Retrieved October 3, 2020, from Nielsen Norman Group: https://www.nngroup.com/articles/research-ops-101/

xvii. Boulton, E. (2019, July 11). The Eight Pillars of User Research. ResearchOps Community (Publication). Retrieved October 3, 2020, from https://medium.com/researchops-community/the-eight-pillars-of-user-research-1bcd2820d75a

xviii. Aguilar, F.J. (1967). Scanning the business environment. New York: Macmillan.
xix. Brand, S. (1999). The Clock of the Long Now: Time and Responsibility. New York: Basic Books, p. 34.

xx. Brand, S., & Saffo, P. (02015, January 27). Pace Layers Thinking. (L. N. Foundation, Interviewer) Retrieved from https://longnow.org/seminars/02015/jan/27/pace-layers-thinking/

xxi. Geertz, C. (1998, October 22). *Deep Hanging Out*. Retrieved October 4, 2020, from The New York Review of Books: https://www.nybooks.com/articles/1998/10/22/deep-hanging-out/

## REFERENCES CITED

Aguilar, F. J. (1967). *Scanning the business environment*. New York: Macmillan.

Boulton, E. (2019, July 11). *The Eight Pillars of User Research*. ResearchOps Community (Publication). Retrieved October 3, 2020, from https://medium.com/researchops-community/the-eight-pillars-of-user-research-1bcd2820d75a

Brand, S. (1999). *The Clock of the Long Now: Time and Responsibility*. New York: Basic Books.

Brand, S., & Saffo, P. (02015, January 27). *Pace Layers Thinking*. (L. N. Foundation, Interviewer) Retrieved from https://longnow.org/seminars/02015/jan/27/pace-layers-thinking/

de la Nuez, A. (2019, August 12). *Democratization of UX Insights: What does this really mean?* UX Matters. Retrieved from https://www.uxmatters.com/mt/archives/2019/08/democratization-of-ux-insights-what-does-this-really-mean

Geertz, C. (1998, October 22). *Deep Hanging Out*. Retrieved October 4, 2020, from The New York Review of Books: https://www.nybooks.com/articles/1998/10/22/deep-hanging-out/

Kaplan, K. (2020, August 16). *ResearchOps 101*. Retrieved October 3, 2020, from Nielsen Norman Group: https://www.nngroup.com/articles/research-ops-101/

Kara, H. (2018). *Democratizing Research in Practice*. In R. Iphofen, & M. Tolich (Eds.), *The SAGE Handbook of Qualitative Research Ethics* (pp. 103-113). ProQuest eBook Central: SAGE Publications.

McKercher, K. A. (2020). *Beyond Sticky Notes*. Sydney: Beyond Sticky Notes.

McKercher, K. A. (2020, October 3). *What is co-design?* Retrieved from Beyond Sticky Notes: https://www.beyondstickynotes.com/what-is-codesign

Reichelt, L. (2019, November 11). *Five dysfunctions of democratised' research. Part 1 - Speed trumps validity*. Medium. Retrieved October 3, 2020, from https://medium.com/@leisa/five-dysfunctions-of-democratised-research-part-1-speed-trumps-validity-1606fa63a3e8

ResearchOps Community. (2019, February). ResearchOps Community: About. Retrieved from ResearchOps Community: https://researchops.community/about

Rosenfeld Media. (2020, October 10). *Advancing Research Program*. Retrieved from Advancing Research 2020: https://rosenfeldmedia.com/advancing-research-2020/program/

Saha Mitra, S. (2020, April 22). *Undemocratising User Research*. UX Collective. Retrieved October 3, 2020, from UX Collective: https://uxdesign.cc/undemocratising-user-research-6897e6f4cf72

Sirjani, B. (2020, July 22). *Democratization is our Job*. Retrieved from Yet Another Studio: https://yetanother.studio/blog/democratization-is-our-job

Towsey, K. (2019, July). *An interview with Kate Towsey, ResearchOps Manager at Atlassian*. (N. Francis, Interviewer) Marvelapp. Retrieved from https://marvelapp.com/blog/kate-towsey-researchops-atlassian/

# Empathy Writ Small & Large

Ethnography as a tool to evoke empathy is often denoted as a good in itself. How might the process of building empathy at scale work for (or even against!) the impact of ethnography in organizations?

Session Curators: Nick Agafonoff, Lindsay Ferris, Tabitha Steager, Erin Taylor

PECHAKUCHA

# How Tragic Flaws Resonate at Scale

## Just Ask Shakespeare

MEGAN DAVIS, *Spendlove and Lamb*

*What does Hamlet, Daenerys Targaryen and Persona development have in common? As an actress I had an in-depth understanding of character development but struggled with creating insights from personas. I embarked on a quest to find out why. I interviewed one of the greatest authority's on Shakespeare, John Bell, AO, OBE FRSN about The Bard's greatest gift to storytelling, the invention of personality and characters. I share the realization that came next, how an individual character's tragic flaw can resonate at scale.*
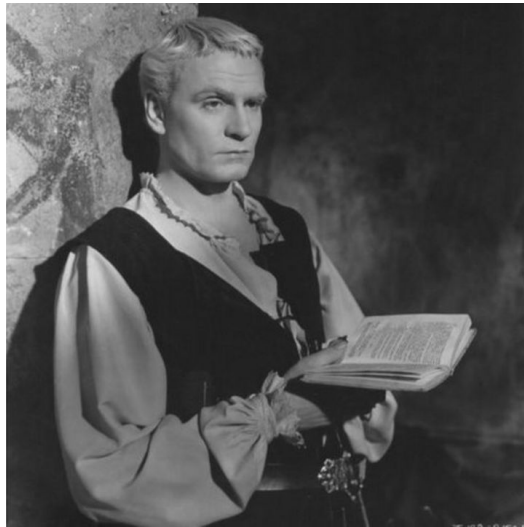
*Hamlet*. Dir. Laurence Olivier. Universal International, 1948. Film.

**Megan Davis** is originally from Michigan and now living and creating in Melbourne, Australia. As the founder and lead storyteller at Spendlove and Lamb, she has been discovering stories since 2012, specializing in narrative based communication strategies, using human centred design practices to create meaningful visions of the future. Each bespoke framework uses stories to strategically move projects, organizations and governments into the new realities they are creating. Megan loves traveling the world physically and virtually to shape narratives for future scenarios, transformation and innovation processes. She has conducted workshops in New York, London, Berlin, and most recently Lisbon at the House of Beautiful Business in 2019 to teach people how to connect with storytelling to transform their worlds. Her latest passion is creating storytelling frameworks for innovation and human-centered design projects to help people connect deeply with their customers. Putting empathy first in business and life, she loves sharing her knowledge by speaking, training, and consulting on projects that are changing the world. megan@spendloveandlamb.com

CASE STUDY

# Harnessing Empathy to Scale a Healthtech Startup During the COVID-19 Pandemic

## A Case Study Of myICUvoice, a Communication Tool Designed for Critical Care

NADYA POHRAN, *University of Cambridge & SympTech Ltd*
TIMOTHY BAKER, *SympTech Ltd*
SIMON PULMAN-JONES, *Emergence Now*
AMY WEATHERUP, *AJM Enterprises*

*This case study explores the scaling experience of an early-stage healthtech startup company called myICUvoice. During the Covid-19 pandemic, myICUvoice rapidly scaled from a single intensive care environment to being widely used nationally (UK) as well as globally. We explore* why *and* how *so many volunteers were motivated to donate their time and expertise to help scale this early stage startup. Specifically, we examine the roles that empathy played throughout the scaling process. There are three distinct types of empathy that we have identified in our story:* em-pathos, *empathetic understanding, and mass-empathy. These each had a distinct role in driving the startup forward. Importantly, we note that human-centered design (which often focuses almost exclusively on achieving empathetic understanding) will immensely benefit from considering the multiple types, and multi-faceted powers, of empathy.*

*Keywords: empathy, pandemic, startup, sustainable innovation, healthcare, volunteers, software*

## GENERAL INTRODUCTION

When it comes to ethnography, the enforced remote-working conditions due to the COVID-19 pandemic have caused many of us who conduct ethnographic fieldwork to contemplate, reconsider, and re-frame the notion of "being there" (Collier Jennings and Denny, 2020; Denny et. al 2020). Ethnography's longstanding commitment to in-person fieldwork has, indeed, been challenged and is being re-explored as a result of global working conditions and what constitutes an appropriate social interaction. As such, it may seem that any paper focused on ethnography and COVID-19 would naturally focus on exploring what it means to (/not) "be there". But our focus in this pandemic-related paper is notably elsewhere—or, at least, we are not directly focused on (/not) being there. Rather, we focus on the way that being *sensitively-attuned to the role that empathy plays* —regardless of whether one is physically *there* alongside others—can be of benefit to the process of scaling a startup, or of innovation more broadly. Suffice it to say, as much as the circumstances of COVID-19 have challenged the ethnographer's notion of being there, the pandemic has *also* created opportunities for the rapid formulation and implementation of ethnography-driven human-centred innovation for certain COVID-19 related startups.

This paper details the particulars of myICUvoice's journey, focusing specifically on how it spent nearly 7 years without any significant scale-up, and then underwent rapid national

and global scale-up during the pandemic. As we detail our own journey, we highlight and explore more broadly relevant questions. We ask, for example: are the conditions of this pandemic providing a new scaling model for startups? Can the experience of COVID-19-related startups be used to inspire new strategies for scaling human-centred innovation that will be of value beyond the pandemic? How can a startup scale when the traditional measures which tend to propel a company forward—investment funds, for example—are absent? What role did multiple stakeholders have in agreeing to be a part of myICUvoice's journey towards scaling? How can volunteers be mobilised in order to allow for rapid scaling? In this paper we explore these questions in a way that resonates with one of anthropology's well-established goals: to use the particular to say something about the general (Tambiah 1985). The particular details of the myICUvoice case study can shed light, more broadly, on some general ways in which ethnographers and startups can pay attention to empathy—we return to some general "lessons learned" in our conclusion. Notably, our own exploration of these above-mentioned questions continually returns us back to empathy; three distinct but complementary types of empathy (we define these below) have been crucial in propelling myICUvoice forward.

In order to explore this annual conference's focus on topic of scale, we offer a model of four phases of scaling that we have identified in the myICUvoice case study. These each present different conditions and opportunities for scaling impact through human-centred innovation. They are:

1. Moving from the defined initial problem to a design; a clinical/technological concept must embrace the complexities of the broader human context in which it can bring impact
2. Moving beyond human value to organisational/institutional value
3. Wide-scale empathy (triggered and amplified by COVID-19) transforms the conditions for scaling human-centred innovations
4. Empathy-driven networks drive a rapid global scaling

The first two of these phases are familiar territory for human-centred design initiatives. But COVID-19 has brought unusual conditions for scaling a startup; whereas the first two phases were driven by individual empathy, the latter two phases in our model hinge upon the "mass-empathy" phenomenon which has been a characteristic of the COVID-19 crisis.

We consider how empathy manifested in three different ways at three different stages. First we consider the ways in which empathy was leveraged by Tim and others from the very beginning stages of this startup. Then, we consider how Nadya's ethnographic work with nurses in the ICU furthered our empathetic understanding of how to reach the startup's objectives. Lastly, we consider how wide-scale empathy of the general public was crucial in the scaling process of our startup; importantly, we consider how this empathy first had to be understood. These three manifestations and stages of empathy are rather distinct, and we define our terms in more detail below.

While we are, of course, fully aware that a global crisis itself is not something that should ever be desired or intentionally replicated,  we wonder whether what this pandemic has enabled us to recognise and identify something that startups could seek to replicate. Empathy, when recognised and harnessed by somebody capable of empathetically-understanding multiple viewpoints, is a very powerful resource. As was evident here, if

"mass-empathy" is to fuel the accelerated commercial growth of a new startup, the innovators would need to identify ways to trigger and leverage it appropriately.

Our mention of "post-COVID" prompts us to explicitly note that we are working with four rough temporal periods in our present discussion: (1) *pre-pandemic* (which we date to anything up until February 2020, since it was around that time that COVID-19 began to be recognised as a state of wide global concern); (2) *acute pandemic* (characterised by an unexpected and unprecedented rapidly spreading global event in which treatment options were limited and individuals' lives were rapidly lost in the midst of overburdened healthcare systems, which we date from March 2020-June 2020); (3) *chronic pandemic* (in which disturbances to health, social order, and economy were a constant threat and a present possibility in the minds of many individuals and governments, which we date from July 2020-ongoing), and *post-pandemic* (when either a cure is available or herd immunity is gained, a time marker which remains unknown.) These four periods are not clear-cut distinctions; rather we view our delineations as a heuristic device meant to enable our present discussion and consequently promote further dialogues and explorations on the topic of startup growth.

Above, we wrote that we focus on three distinct but complementary iterations of empathy. (1) There is the kind of empathetic understanding ("*verstehen*," to invoke the German term common to anthropological discourses) in which an individual seeks to gain an understanding of another's viewpoint. In human-centered design conversations, this empathy is often spoken of as being a crucial step in the design process: one has to figuratively step into someone else's shoes if one wants to design an effective and useful product for them. This iteration of empathy can indeed be heartfelt and have an emotional component to it, but, in the case of human-centered design, it is notably task-driven and goal-oriented: the reason for attaining this empathetic understanding is to design a viable and useful product.

(2) The second iteration of empathy that we explore is one that is likely more familiar to the way the word is commonly-understood: it is the uncontrollable surge of emotions—often connected to heartache, loss, grief, sadness, etc. —which prompts one to imaginatively experience what it is like to live, breathe, and operate in the world from another perspective; consequently, one feels an empathetically driven desire to improve that person's situation. While this sounds, at first glance, similar to the first iteration of empathy detailed above, this second iteration of empathy hinges upon the emotions and deeply held feelings. Etymologically, the word's Greek origins are *em* (in) and *pathos* (feeling), and they suggest a sort of *feeling with/alongside/for* someone else, especially when that person is struggling or suffering. While, at times, an empathy for the suffering of others can in fact cause us to stand still in our tracks as our brains try to process the seemingly overwhelming obstacles involved, empathy often compels us to some degree of compassionate action. That is, when we empathetically-resonate with the sufferings of others, many of us are inclined to do *something* to help.

'Empathy', then, can be *either* strategic or spontaneous; premeditated or unprompted; it can be a skillset which is trained and fine-tuned for optimal execution and maximum efficacy, or it can be a surge of unanticipated emotions which render one's heart so full that it cannot help but begin to spill over. In this paper, we will distinguish between these first two types as *empathetic understanding* and *em-pathos*, respectively.

(3) We also refer to a third type, which we call 'mass-empathy', and by which we mean a large-scale, collective experience of *em-pathos* for the same phenomenon. In this particular

case, it was the COVID-19 pandemic which fueled mass-empathy for ICU patients, and for healthcare more broadly.

For their part, ethnographers have often spent years training in empathetic understanding and it is often so deeply embedded in their ways of seeing the world that they cannot help but adopt and actively cultivate a posture of empathetic understanding even outside their official role as ethnographer. It is interesting to note, though, that applied ethnography and the anthropological tradition from which it emerged has a strong bias towards rational, reflexive *empathetic understanding.* This paper explores what can be gained from opening up our awareness of the role of *em-pathos,* the dimension of empathy we might otherwise suppress or exclude.

In myICUvoice's journey, we experienced multiple distinct and sometimes overlapping layers of *em-pathos* and empathetic understanding. As individuals, we are driven towards ever-deepening our own abilities to draw upon both of these types of empathy as we interact with the world. Indeed, as we hope will remain clear throughout this paper, we firmly believe that these kinds of empathy are key in building a kinder world. Regardless of these personal convictions, our present exploration is focused on scaling and, thus, our crucial point is this: individuals who are tuned-in to recognise, cultivate, and harness the *em-pathos* that compels individuals towards generous action in themselves and others not only offer a helpful vantage point and sounding board in what we could broadly define as business endeavours, but they are in fact *also* vitally-placed to drive forward projects which have historically struggled to pass through the inevitable obstacles involved in scaling. That is, an individual who has the ability to empathetically-understand and wield *em-pathos* itself can be wondrously helpful in driving startups. Because, as we will show, *em-pathos* can be a powerful driving force which, when properly harnessed, can pull along a startup into a rapid scaling process, it is important for individuals to be able to recognise and successfully harness it.

While it is not necessary for such individuals to be anthropologically-trained, it is this ability to focus on and contextualise small and seemingly-mundane details—that is, the sort of stuff greatly emphasised by cultural anthropologists and others who conduct long-term ethnographic fieldwork—that can powerfully drive forward a startup's vision, and enable it to scale even when the features which have traditionally helped startups to scale (new defensible IP, financial investors, etc.) are insufficient.

## INTRODUCTION TO myICUvoice AND CRITICAL CARE

Before returning to articulate precise moments of the four-part model that we offer for scaling, let us start with an introduction to the case study itself: a Cambridge (UK) based startup (SympTech) developed a product called myICUvoice. We will walk you through the journey of this particular startup, highlighting both the types of hurdles it encountered as a software startup in a healthcare setting, as well as its engagement with an ethnographer and human-centered design and the way that this collaboration, at various times, either challenged, changed, or strengthened the startup's visions and processes. Additionally, we will offer specific and detailed examples of the role that empathy played throughout the process.

So, what is myICUvoice? One way to succinctly summarise myICUvoice is that it is a specially-designed communication tool for mechanically-ventilated ICU patients and their nurses in the form of an iPad app. Mechanically-ventilated patients cannot use their vocal

chords to speak, and this inability—combined with patients' frail physical states and other factors which inhibit more standard forms of communication—impairs the patients' ability to communicate with their clinicians and family members. This lack of communication, of course, also means that clinicians are not able to fully understand their patients' wants and needs. Research shows that, while all patients in the ICU naturally experience some fear, stress, and uncertainty during their time in critical care, over 25% of patients who leave the ICU experience some degree of anxiety, post-traumatic stress disorder, and depression (Wade et. al 2018).

The driving energy behind the creation of myICUvoice can be simplified to a set of linked presuppositions: (1) Patients, like all individuals, need to communicate about their basic needs and wants as well as the types of interactions (humour, gratitude, a request for more information, etc.) which improve their overall quality of life. (2) Due to inadequate communication, the suffering that many patients experience during critical illness goes unrecognised, and hence untreated. (3) An inability to communicate has other long-lasting undesirable effects for patients. (Even though explicit recall of these events is often compromised by illness and drugs, patients report these problems after recovery from critical illness. In reality, such symptoms are likely to be pervasive and the implicit memories they encode may contribute to the increasingly recognised problems of acute post-ICU delirium and late psychological problems such as post-traumatic stress disorder.) (4) When communication is restored, the ICU experience is better for all involved; the earlier that communication is restored for patients, the less severe their PTSD will be upon leaving the hospital environment.

## TIMELINE OF THE myICUvoice CASE STUDY

With that overview in mind, let us move to the timeline of myICUvoice. myICUvoice as a communication app was first conceived in 2013 by Dr Timothy Baker, a specialist registrar in anaesthesia and critical care medicine, while he was working in the ICU of Addenbrookes Hospital, a major teaching hospital in the UK. Along with his ward boss Dr. Vilas Navapurkar, Tim started a patient focus group—the first of its kind in the UK—in which seven former ICU patients were invited to share what it was like to be a patient in the ICU.1 The doctors had realised that, although their ICU was one of the most successful in the country in terms of medical metrics (namely, their standardised morbidity rate was notably low), they could likely still be doing something better, and they wanted to learn from their patients what this something was. We can see empathetic understanding within this desire to understand the viewpoint of others.

This patient focus group revealed to Tim and Vilas that, while patients rationally understood that they had been cared for by talented doctors, and while they often possessed an immense gratitude for the doctors who had managed to keep them alive, they also did not feel that they had been treated (in both the medical and non-medical senses of the term) as individuals. They were, first and foremost, patients. And an intrinsic part of being an ICU patient, it seemed to them, was to not have a voice. A clear spectrum of voicelessness was described; some patients were in the ICU ventilated whilst awake via an artificial airway (physically preventing speech) for weeks or even months, others were too weak to have the cerebral ability to find the words or physically phonate, others felt oxygen masks prohibited being heard and some felt that the complexity of critical illness meant they had no say in

their treatments. These experiences of voicelessness also extended to not being able to express even their simplest needs such as "I am thirsty, can I have another drink please." Not feeling able to fully communicate led many patients to ultimately conclude that the doctors really did not understand their experiences (see also Alasad et al 2015; Lykkehaard and Delmar 2015; Moen and Nåden 2015; Samuelson 2011; Topçu et al 2017).

Through the 2013 focus group, we see both the effort to gain an empathetic understanding of the patients' experiences, and, upon gaining this, we see a deep sense of em-pathos emerge: Tim concluded that something must be done in order to better understand the thoughts and feelings of his patients and, crucially, to enable them to have a voice while in the ICU. Moving from empathetic understanding to em-pathos, Tim felt compelled to create a communication tool for patients as a way of restoring their voice. He wanted them to be able to communicate their physical states, but also to be able to express other elements which are basic to human interaction: humour, gratitude, questions, and the like. In other words, this empathetic (em-pathos) drive planted the seed for myICUvoice. Because startup innovations come out of the solvable problems, rather than the intractable ones, recognising a solvable problem must necessarily predate any efforts to scale that solution.

## Phase One: Moving From Initial Problem to Design

In 2014, because of a lack of available funding, Tim approached his cousin (a student with no prior software development skills), who, from a place of *em-pathos*, agreed to develop a simple prototype to be trialled in the Addenbrooke's ICU. This cousin then relayed the story of myICUvoice to a graphic designer, highlighting, as Tim had done, the reality of mechanically-ventilated patients and the severe effects of not being able to communicate. The graphic designer responded with *em-pathos* and created tiles, thus providing images to match the app's text statements. Version 1.0 of myICUvoice was created: with the press of a button, patients could use the purposefully-simplistic tiles to select from pre-written statements to express communications about their physical and emotional needs, and also ask questions ("where am I?", "when is my family coming back?") that are typical to ICU contexts.
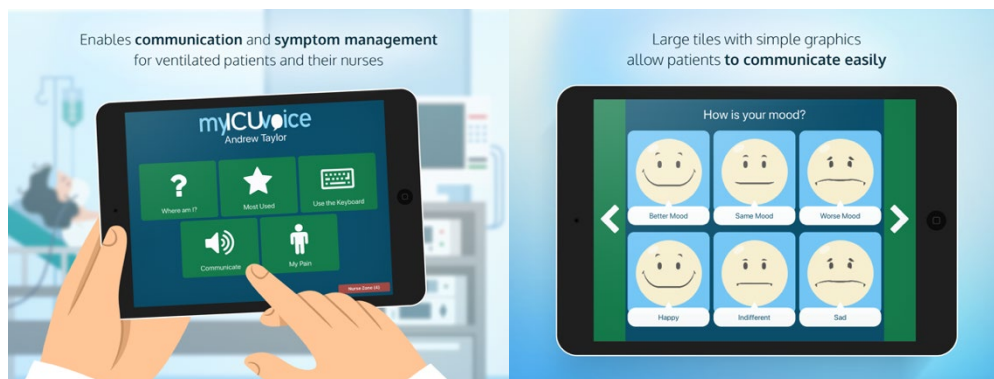


Figure 1. An image featuring the myICUvoice homepage (left) and an image of the "How is Your Mood?" page of the myICUvoice app. © SympTech, used with permission

A study comparing patients with and without the provision of the tool demonstrated enormous benefit for patients. Nurses, doctors, patients and their relatives provided a continuous stream of positive feedback. Despite this, patients that were deemed clinically suitable for the tool were often found not to have been given access to it. Some explanations offered for limited provision included nurses not realising the patient would be able to communicate, iPads not being available, having been lost or without charge and the fact that it was clearly a prototype rather than a polished solution.

.

**Percentage of the patient and staff groups who strongly agreed or agreed with each survey question**

| Response Group | | Patients | | | Staff |
|---|---|---|---|---|---|
| Question | | Control | Conventional Communication | myICUvoice | myICUvoice |
| Patients often felt | Stressed | 40.4% | 83.5% | 71% | 51.3% |
| | frightened | 36.5% | 83.5% | 71% | 32.5% |
| Staff Understood | Physical needs | 75% | 83.3% | 100% | 87% |
| | Emotional needs | 67.3% | 33.3% | 85.7% | 71% |
| | Levels of discomfort | 63.5% | 50% | 85.7% | 85.9% |
| Communication was effective | | 41% | 33.3% | 57.2% | 87% |

Figure 2. Numerical data showing the percentage of the patient and staff groups who 'strongly agreed' or 'agreed' with each survey question (Baker et. al 2017). Reproduced with permission.

This initial work was used to demonstrate feasibility of the project and to highlight the technical areas that needed to be improved or established. In 2016 a grant for development was provided by Addenbrooke's Charitable Trust (ACT) to build a prototype that included additional features. This was introduced onto the ICU in 2017 and included a database to catalogue the symptoms reported and the extent of use, and measure usage patterns. Patients who had the appropriate dexterity and muscle control could also use a new keyboard function that supported free-communication allowing expression of anything desired. Despite the technical enhancements in the software, there was no improvement in the rate of use of the application in the Addenbrooke's ICU.

We have retrospectively suspected that this lack of uptake is directly linked, as we will return to below, to the fact that myICUvoice had been designed with ICU patients in mind, but it was in fact ICU nurses whose role as end-users needed to also be considered. Up until this point in the myICUvoice journey, it was ICU patients who were the key focus as the end-user whose needs had to be met and whose preferences had to be catered to.

But, by the nature of healthcare and the subsequent interpersonal relations that occur in critical care contexts, the patient was in fact a vulnerable dependent who relied on the nurse to introduce, and continue to use, the myICUvoice tool with them. Like so many other

communication devices, we had designed our tool for one set of end-users while there were in fact multiple end-users whose needs had to be considered if we wanted to stimulate uptake. Indeed, healthcare startups must understand the dynamics and decision-making processes of the healthcare systems (which differ by country) and ensure that they can explain tangible benefits to the gatekeepers while also offering a helpful therapeutic solution to patients—we return to this in more detail when we explore the role that ethnography and human-centered design played in the journey.

By 2018, several presentations given by Tim and colleagues piqued the interest of some charitable donors, including more donations from ACT, which resulted in accruing additional iPads so that myICUvoice was available to all Addenbrooke's ICU patients 24 hours a day. In addition to the provision of hardware, to ensure that the tool was always available for the physically voiceless patients but also so that a better understanding of the symptoms of all patients could be gained, myICUvoice was used to survey all patients in the ICU on a daily basis. Nurses were trained to use the system, and the rationale for allowing improved communication was explained. At this point in time, the main methods of training nurses included emails (reminding them to use the tool with every patient and with a link to an explanatory video of why the tool should be used), study day presentations, and peer to peer learning (five dedicated nurses were first given additional training to teach and encourage other nurses of the same grades.)

Interestingly (or, frustratingly, depending on your own experiences with this sort of thing) although Tim and colleagues felt that they had clearly explained to all nurses how, when, and why to use the myICUvoice prototype, ethnographic observation and interviews conducted in late 2018 suggested that very few nurses felt that any of these topics had indeed been explained to them. Rather, many expressed a desire to receive more training about how to use the device. A number of nurses, even those who had been on the receiving end of educational material about some of the broader reasons for using myICUvoice, still had not fully comprehended what Tim and others felt they had made crystal clear. This striking incongruity between what (on the side of the startup) was thought to be clear and what (on the side of one group of end users—the nurses) was felt to be clear is important to note. Indeed, one of the struggles with startups that hinge upon creating and providing educational training to any group of their end-users is that one cannot always be certain that the training has been adequately received. Training must be designed and delivered in a way that is relevant to its end users and which highlights their areas of interest. (Again, we return to this below when myICUvoice's timeline interacts with ethnography.)

As the prototype remained in use in Addenbrooke's ICU, nurses were encouraged to incorporate it into their standard care routines and to use it with every patient at least one time each day. Despite positive feedback received from the medical staff and patients who used it, the data that the app produces demonstrated that it was not being used with all patients who would benefit from it, and some nurses rarely used it at all.

myICUvoice wanted to use data visualisation tools as a way of addressing this apparent gap between who could be using it and who was using it. Consequently, Phil and Mary-Ann Claridge, of Mandrel Systems (a software consultancy firm), were introduced to the project: they were told about the need for mechanically-ventilated patients to communicate, and they wanted to offer their time and expertise to help the myICUvoice project. Em-pathos was central here: having seen his own father undergo medical care which required a tracheostomy, Phil directly resonated with the aim to provide better communication for

ventilated patients. On the other hand, Mary-Ann had a family connection to Florence Nightingale (a statistician and founder of modern nursing), and she was keen to be a part of using data to improve patient experiences. Consequently, Mandrel Systems volunteered their time in order to design and create the necessary data visualisation.

The live data visualisation tool allowed for further exploration of myICUvoice's efficacy. It demonstrated symptom patterns not previously recognised and new treatment strategies to be trialled. It also meant that usage data could be collected on a daily basis. When we were able to review the data collected by myICUvoice every day, it became clear that the level of usage reported by the nurses didn't align to the data extracted. It also suggested that the tool wasn't being used in the way we had intended. What we found was that only on days when nurses were actively and repeatedly asked to ask their patients to use myICUvoice did we see a level of use close to that intended. This was when Tim first started to see fully the impact of the views of the nurses in determining whether the software was being used appropriately.

Realising the need for sustainable funding and investment in the technology to make a Minimum Viable Product (MVP), Tim engaged 8 MBA candidates from the Judge Business School (University of Cambridge) in 2017. The question asked was, 'We know myICUvoice creates value, but making patients feel better doesn't save money for hospitals. How can we make a viable business model for myICUvoice?' They conducted market research, demonstrated the clear need for the tool, understood the value proposition but found no route to market without a multi-centre randomised control-trial to demonstrate benefit and therefore cost benefit to healthcare (this is the classic medical business model) This is to say: myICUvoice faced major hurdles to achieve buy-in from hospital and wider national health service authorities. The resounding conclusions of the MBA candidates was that there was no compelling way for myICUvoice to proceed as a regular startup (i.e. one that is financially self-sustaining, let alone profitable) and it would only become operable on charitable funds.

While the progress thus far had been a long labour of love by Tim and other individuals who had joined his em-pathos driven vision for the app, the other three authors of this paper were each introduced to myICUvoice in 2018, when it participated in a Cambridge University technology commercialisation program called i-Teams (Pulman-Jones and Weatherup 2019).3 Nadya Pohran, who at the time was a postgraduate student participant of i-Teams working specifically on the myICUvoice project; Amy Weatherup, the founder and director of the i-Teams program; and Simon Pulman-Jones, an instructor and mentor on the program. The i-Teams program, which was presented at EPIC 2019 as a model for the human-centred design community to engage earlier with scientist innovators, worked with myICUvoice in early 2018. i-Teams takes teams of post-graduate scientists through a kind of participant-ethnography simulation of the startup experience, with a strong human-centred design ethos, with the aim of supporting and accelerating the commercialisation path of new university-developed innovations. The student i-Team included 4 life scientists working towards their PhDs, a Chemist studying for a Masters degree, a postdoctoral clinical scientist and Nadya, who was working towards her PhD in Theology and Religious studies while using anthropological methods. During the project the team talked directly with doctors and nurses who worked with ICU patients, and to some patients' families, to increase the empathic understanding of the benefits that myICUvoice could bring to them. They also investigated possible sources of funding and routes to market for the application. The project highlighted the potential for improving the value and relevance of myICUvoice through a deeper understanding of the different stakeholders in the ICU experience. The

team recommended forming a relationship with another ICU to try the application there, and then investigating whether it would be possible to gather data to demonstrate its effectiveness in an objective way, through some type of formal trial.

When the i-Teams program had finished, Nadya continued working with myICUvoice over a 14-month period, deploying elements of ethnography and human-centered design to inform a significant redesign of the myICUvoice app. Nadya conducted several half-days of observation in the ICU where she was able to interact with, and observe, nurses as they cared for their patients. This totalled to ~20 hours of observation. She also conducted semi-structured interviews with 25 nurses, and she was undeniably interested in finding out why some nurses were not using the app, as well as gaining a more general understanding of how the app was perceived by the nurses. However, she made a point of not posing her questions directly to nurses, but instead she mentioned her affiliation with the myICUvoice project by means of introducing herself, and then went on to explain that she simply wanted to understand more about nurse and patient interactions in the ICU more broadly. This choice of intentionally allowing the nurses to direct the subsequent topics of conversation resulted in her assuming the posture of a student who nurses often seemed to view as someone in need of teaching and training. This methodological posture of "open-student", certainly familiar to Nadya's background in cultural anthropology, is also familiar to medical settings, where nurses (depending on their expertise and training) are often in the position of training more junior nurses or medical students. All this to say, it was a methodological posture which worked strikingly well in the ICU setting.

Typically, Nadya would arrive at the ICU, put on scrubs, and enter one of the three wards of the ICU, allthewhile wearing her hospital ID card which stated "Ethnographer and Usability Auditor" as her job title. (Many people didn't read past the word "Ethnographer" before asking something along the lines of 'Ethnographer? What's that?!') In ICU contexts, each nurse is assigned to one specific patient during their shift, and so, while it was relatively easy to locate nurses to (hopefully) speak with, it took a certain combination of courage and luck to approach a nurse who deemed that they had time and willingness to speak. Some of the nurses had been introduced to Nadya by Tim or the Matron during her first couple of times in the ICU, but many of the nurses had no idea who she was until she approached them with her line, "Hi, I'm Nadya. I'm working on the myICUvoice project with Tim. Right now I'm just trying to learn more about the ins and outs of the ICU, do you have any time to chat with me?"

While there were several varied human-centered insights that we do not have the scope to explore here, there are a couple that are worth noting, as they have had a direct impact on the technology of myICUvoice. Midway through her time conducting ethnographic observation in the ICU, Nadya began to notice the varied ways that nurses, either directly or indirectly, offered explanations pertaining to why they did not use the myICUvoice app as often as it might have been used. One crucial insight that was eventually brought to light was that nurses often thought of a very particular "ideal" patient with whom they would use the app, and they were not naturally inclined to use it with patients who did not have those particular characteristics.

This idea of an "ideal" patient with whom to use myICUvoice was significant and seemed ubiquitously held amongst nurses, even those who had shown themselves to be somewhat of a myICUvoice ambassador with their vigilant and keen use of the app. Of the 25 nurses who Nadya interviewed, 22 of them either explicitly (by which we mean some

used the word itself) or implicitly referred to an "ideal" patient with whom to use myICUvoice. The 3 nurses who did not do this were ones who self-described as not using the app at all.

This nurse-held belief that there was a particular kind of patient with whom they should use the app, and thus, many other cohorts of patient who did not need to use the app, is particularly significant because, for months, Tim had been instructing nurses to use myICUvoice "with every patient, every day." The "ideal" patient who the nurses conceptualised had defining characteristics: was understood to be mechanically-ventilated, a "slow wean" (that is, that they were expected to spend substantial more time in the ICU before being discharged) and sufficiently alert/awake/oriented (described in nurse terms as Level 2 or Level 3.)

Similar views were expressed when, months after the interviews, Nadya conducted an anonymous survey with the nursing staff. She asked the question "Below is a list of words which could describe a random patient. Please select all that you would try to use myICUvoice with." The words were: irritated, sedated, semi-conscious, tracheostomy, mechanically-ventilated, lonely, confused, happy, can speak verbally, nervous or anxious, Level 1, Level 2, Level 3, spaced out, elderly, middle-aged, young, hallucinating, ward-able. Nurses were instructed to check off all adjectives that they would be willing to try to use myICUvoice with, and a space was given for them to add any other adjectives or comments. Of the 40+ nurses who filled out the survey, 80.6% selected "Level 3", 90.3% selected "mechanically ventilated", 90.3% selected "Level 2", and 96.7% selected "Tracheostomy"— these kinds of answers were to be expected. But, significantly, only 41.6% indicated that they would try to use myICUvoice with a patient who could speak verbally. The fact that less than half of nurses were inclined to use myICUvoice with a patient who could speak verbally is a stark contrast to the way that Tim wanted myICUvoice to be used with all patients. As mentioned above, Tim understood that the communication needs of all patients were not being adequately met and thus, even for those who could speak, communication had to be improved. The fact that nurses were only inclined to use it with a certain patient cohort meant that a mere fraction—about 10%—of patients who could be benefiting from myICUvoice were being introduced to the tool.

This ethnographic insight caused us to reconsider the reality that different end users would inevitably have different reasons for wanting to use the app or not. While it can be tempting to present an apparently well-thought through technical solution and assume that it will be seamlessly integrated and implemented by the various individuals who need to use it, ethnography shows time and time again that the process is not nearly so clean-cut. The reality, certainly in the context of healthcare, is that there is not a single end-user. Consequently, designing the product with the patient in mind (the result of both empathetic-understanding and em-pathos) needs to also consider the needs of the other end users. In the case of myICUvoice, we came to identify several different primary end-users, including: patients, nurses, health-care assistants, doctors, and the relatives of patients. We also knew that our software had to also speak to the priorities of those in positions of power who made decisions about whether or not to implement a particular tool across the hospital setting, and the medical researchers who could potentially use the data contained in the app (we return to this below).

This reality was combined with the fact that the very category of "nurses" (easily spoken of as if the shared title and training necessarily entailed a monolithic and homogenous group

of individuals) quickly proved to be not singular but multiple. To use human-centered design vocabulary, there were multiple personas of nurses: the level of initiative that nurses took in order to excel in their work environment, the intensity with which they wanted to understand the needs and desires of their patients, the extent to which they believed that communication would indeed improve ICU experiences, and their comfort with new tools and technologies on which they had not been extensively trained, are just a few examples of the various defining of features of nurses.

So, while some nurses did not use myICUvoice due to feeling that they did not have an appropriate patient with which to use it, others had different reasons for not using it. Some were not comfortable with the technology. Some seemed generally disinterested in adding more work into their days. Others preferred using alternative modes of communication, feeling that the technology was an inhibitory blockage in giving their patients the human-human interaction that they needed. As one nurse explained, "As soon as the patient can communicate otherwise, I prefer to stop using the iPad [myICUvoice]—[I instead want] to have them mouth words, to nod their head yes or no, whatever. It feels more like a human interactive than using technology…we need human elements of interaction."

Another nurse, whom Nadya had interacted with on many occasions, including seeing him introduce myICUvoice to several patients who had never used it before and who overall seemed immensely on board with the values and mission of myICUvoice, often used alternative modes of communication when possible. One day, when his tracheostomised patient began to convey (through non-verbal means) the early signs of wanting something, this nurse could have easily used myICUvoice in order to figure out precisely what the patient wanted. Instead, the nurse chose to engage with the patient without using myICUvoice, and eventually learned that what the man wanted was a drink of water. (Something which, as it turns out, is one of the most frequently used statements communicated using myICUvoice.) The nurse quickly got water for the man, and the communication mystery was deemed settled. What is striking is that, in this instance, when Nadya eventually had the opportunity to ask the nurse about why he did not use myICUvoice in that instance, he simply shrugged and explained "he didn't need it."

This succinct statement is actually of utmost significance for human-centered design contexts. Who determines when a communication tool is needed or not needed? What if, as strongly suspected by Tim, using myICUvoice in such instances might have in fact lead to more in-depth communications? What if a patient, even when their need for thirst is met, has more complex—or even simple—needs which would benefit from the opportunity to communicate with more nuance and detail? What if, instead of waiting until 'normal' communication failed and using myICUvoice as a 'last resort' technical solution, it was used as the first and consistent means of communication throughout a patient's stay, as Tim originally envisioned? Such considerations are ubiquitous and crucial in the context of human-centered design and especially in the implementation of new products in environments with existing protocols and practices.

## Phase Two: Moving beyond Human Value to Organisational/Institutional Value

When Nadya spoke with Tim about the insights gained from observations such as these, she emphasised the need to clearly define the multiple end users and to recognise that each

set of end users would have different reasons for using the product (or not using it) and that, even within a single set of 'end users' there would be significant diversity in that user group. She noted, for example, that even the end-user of 'patient', which was already diversified on account of different patient cohorts, was further diversified by whether or not we were considering the needs of *current* patients (in which communication was an obvious remedy) or *future* patients (in which the software's ability to track and analyse frequently-reported symptoms, thereby enabling the possible pre-emptive treatment of symptoms for future patients, was most desirable). Furthermore, as mentioned above, the 'nurses' user group was particularly significant to the success of this app, and it was determined that the wide spectrum of 'nurses' had to be empathetically understood and their needs had to thus be accounted for and addressed.

These insights called for significant redesigning of the myICUvoice project: both the app's software, but also the way it was presented to nurses (or, to these varying personas of nurses) had to be accounted for. It was at this point that myICUvoice concretely realised an aspect that, as we mentioned above, had previously been unarticulated but was still very significant to the app's (lack of) scaling process: its end-users were not simply ICU patients; they were *also* ICU nurses.

## Different uses, different users, different priorities

| | Health Care Assistant | Nurse | Consultant | Patient | Family Members | Funders | NHS/ implementation | Medical Research |
|---|---|---|---|---|---|---|---|---|
| Communicate practical needs | ✓ | ✓ | | ✓ | | | | ✓ |
| Communicate pain | ✓ | ✓ | ✓ | ✓ | | | | ✓ |
| Communicate emotions | ✓ | ✓ | | ✓ | ✓ | If linked to improving patient experience | | ✓ |
| Communicate (no set purpose) | | | | ✓ | ✓ | If linked to lessening PTSD | If linked to lessening PTSD | If linked to lessening PTSD |
| Speed up communication | ✓ | ✓ | | ✓ | ✓ | ✓ | ✓ | |
| Decrease PICS and PICS-f | | | | ✓ | ✓ | ✓ | ✓ | ✓ |
| Connect symptoms with conditions | | | | | | ✓ | ✓ | ✓ |

Figure 3. A chart showing the ways that different users were understood to have different reasons to use the app. © SympTech, used with permission.

On the surface level, this can seem like a shockingly simple recognition: the sheer diversity of human experiences and the multiplicity of end users meant that a tidy technical solution, even if it were a perfectly functional as a technical solution to a known problem, risked being discarded by a significant number of our target users if their own user needs were not anticipated and accounted for. This recognition, which was nuanced and developed from numerous back and forth conversations between Nadya and Tim and other individuals who represent a host of disciplines, led to the need to re-brand the app in a way that ensured that nurses were also viewed and accounted for as end-users. This recognition from the

ethnographic work eventually planted the seed for a substantive change to the software: myICUvoice had to be re-designed in order to cater to *nurses* just so much as it catered to *patients*. Unfortunately, a lack of funding meant that this change could not be immediately implemented.

Up until this point, the app's interface had only been explicitly designed for patient use; although we wanted nurses to use it and even had features within it that were there specifically for nurses, we had not made any tangible indicators that the app itself was designed for nurse use; everything was 'the Patient Zone' and thus myICUvoice, in many ways, seemed at risk of joining the fate of all other communication apps that had been designed for ICU patients: they were used by a small handful of nurses who were keen enough to want to use them, but, on the whole, they never became an integral part of the care plan or the healthcare system, and they were consequently never consistently and widely used by patients.

A lack of financial resources did not allow the team to substantially pivot the app in order to address this crucial insight. Despite interest being shown from others in healthcare settings, a consistent lack of funding prevented it from picking up the momentum it needed in order to scale larger than the ICU it first started in. Some software glitches eventually resulted in the app ceasing to function for about a 6-month period towards the end of 2019 and early 2020. Those of us who are familiar with the success rates of startups (roughly 1 in 7 startups do not continue beyond their second year) will not find it altogether surprising that a startup—even one which offered a relatively good solution to a relatively notable and prevalent social problem—struggled to scale.

## Phase Three: "Mass-Empathy" (triggered and amplified by COVID-19) Transforms The Conditions For Scaling Human-Centred Innovations

What is, however, surprising—and what makes this case of particular interest to our present discussion of scaling—is the sheer rapidity with which myICUvoice began to scale in late March 2020. If, in late 2019, myICUvoice seemed to be joining the metaphorical graveyard of communication apps, in March 2020 we were rapidly resurrected. Significantly, this was mere weeks after the COVID-19 virus was declared a global pandemic by the World Health Organization, and the fact that myICUvoice's speed of scaling corresponded to increasing global media coverage about the pandemic and ICUs in particular is a crucial aspect for our consideration (World Health Organization, 2020.) While the *em-pathos* of several particular local individuals had driven the app forward thus far, the startup was now operating in a global context of 'mass-empathy.' Making use of the global *em-pathos,* by June 2020, myICUvoice underwent substantial changes and made substantial technical and market adoption progress in a mere two-month period. Having been in the works for nearly 7 years, this shift is notable.

In the middle of March, Tim contacted Nadya to see if she would be willing to act as Project Manager of myICUvoice as it resurrected itself and prepared for scaling; having witnessed her ethnographically-inclined ways of engaging with myICUvoice thus far, Tim sensed that she intuitively 'got' the human-centered vision that he had, and that she would bring a depth and helpful vantage point to the process on account of her anthropological background. For her part, given that unrelated international ethnographic fieldwork had been unexpectedly cut short due to the pandemic, and sensing that it would feel rewarding

and uplifting to be a part of a project that promoted positive change in the midst of the chaos of the pandemic, she agreed to take on the role. Tim had already been joined by Phil and Mary-Ann Claridge who, yet again, crucially offered their time as software developers, as well as by Dr. Katy Surman who offered her time as a medical researcher and general assistant to the various tasks that required attention. These five individuals formed the 'core team' of myICUvoice, as the company (which didn't even yet have a bank account) continued to have a working budget of $0.00.

All individuals were working pro-bono, but they held weekly group meetings and exchanged various emails and messages each day in order to ensure that the tasks were being attended to. The hours were long, and the team furthermore had to work around Tim's own schedule as a doctor attending to COVID-19 patients in the ICU. While this made scheduling challenging, it also meant that we had regular continual feedback from the rapidly changing situation when it came to hospital protocols related to COVID-19. Additionally, because COVID-19 meant that ICUs across many different countries were receiving a significant influx of patients who had to be mechanically-ventilated, it was clear that myICUvoice would be of real benefit to these patients if it could be made available. First and foremost, the software itself was resurrected and certain technical features were redesigned. Specifically, the aforementioned need to design the app for nurses was finally addressed, and the team designed, trialed, and re-designed a 'Nurse Zone' which was specifically designed to provide nurses with impetus for using the app in a way that had not previously existed. While the app's interface had been thus far designed for patients (we call it the Patient Zone), the Nurse Zone had features that were specifically designed for Nurses. For example, it contained summary charts of their patients' symptoms, and showed in a single glance how their symptoms and communications were changing over time.

It is important at this point to note that, while the core myICUvoice team was working (sometimes more than full-time hours) on this project out of their own *em-pathos* and desire to launch myICUvoice to the patients who would, they felt, immediately benefit from it, this in and of itself was not an altogether new scenario. As the previous sections have detailed, there had been a plethora of individuals who had voluntarily donated their time and expertise to the project over the past several years.

But an obvious difference between those 6-7 years and these 2-3 months were the wider global context in which those 2-3 months occurred. The COVID-19 pandemic resulted in conditions that had not yet been previously experienced. As COVID-19 pushed ICU patients and patient experiences into the public eye, more attention of the public was directed towards the ICU. There was an urgent need to optimise efficient use of scarce ICU resources through and and all means, as well as a need to support relationships with patients' families and carers remotely. Meanwhile, because the media coverage focused on the predicament of patients and health professionals (starting in Wuhan, then in northern Italy, and then in other milieus), individuals were being confronted with the *empathetic understanding* of the ICU experience in a way that many had not previously done; for many individuals, *em-pathos* ensued. With its human-centered, nurse and patient-focused design, myICUvoice was well-positioned to respond when conditions changed and new opportunities for scaling up arrived, allowing the team to take advantage of the new ground-swell of mass-empathy for ICU patients which the pandemic generated. It was at this point in myICUvoice's journey that *em-pathos*, empathetic understanding, and mass empathy collided.

**Phase Four: An Empathetic Understanding Of Empathy-Driven Networks Drive A Rapid Global Scaling**

Just as a gust of wind will not propel a sailboat if the sails are not hoisted and ready, a time-sensitive global phenomenon of mass-empathy will do nothing to propel a startup to scale if the winds of *em-pathos* are not adequately harnessed. Empathy, as we have already outlined above, is often spoken of in one of two contexts: either the empathetic understanding of another's viewpoint or the heartfelt and often spontaneous *em-pathos* which compels individuals to act towards bettering another's circumstances. At this point of the four-phase model of scaling, we relied on an empathetic understanding of *em-pathos* itself, and we subsequently made use of the significant expressions of *em-pathos* felt by the broader society in order to drive forward our scaling.

How precisely did this happen? The story is woven together by a few distinct threads. In late March 2020, on day six of her quarantine following international travel, Nadya decided to launch an ethnographic diary study in order to learn more about individuals' COVID-19 experiences. Driven by a seemingly insatiable desire to empathetically-understand the world around her, she felt that a diary study might lead to interesting insights about human thought and behaviour. She wondered, broadly, what sorts of thoughts and actions were new, prevalent, or falling to the wayside during these unprecedented times. As she began to design the study and create the questions themselves, she contacted a handful of colleagues, some of whom worked in policy implementation and mental health fields, with the aim to see if there could be any way to use any insights that the diary study might bring to light as means of invoking some degree of positive change. Colleagues advised her that it was not likely that such a study (detached from any university or company) could be used to change any policies, but that perhaps she could find some insights within the study and eventually implement micro-changes at local levels.

By this time, the diary study had attracted ~100 participants, and Nadya was keen to see what might come of it, so she continued all the while knowing that 'nothing official' might ever come of it. Participants, who were first briefed about the point and purpose of these diaries, provided some basic sociodemographic data about themselves and their living situation and then, over the subsequent three-week period, they wrote an anonymous diary every day. The diaries followed a similar structure each day: participants were first asked to indicate the types of emotions that they had experienced throughout that day, and to provide a descriptive summary of their day including 3 specific details or activities. They were then given three guided questions to reflect upon and answer—the questions had been mapped out in advance to cover a range of thematic topics. And finally, they were given a blank space to add anything else that they wanted. As she read through the entries, Nadya was struck by how many individuals expressed a desire to do something—anything—to feel both useful for, and connected to, the wider world. Much to her surprise, many participants indicated that they appreciated even participating in the diary study itself, as it made them feel as if they were being connected to something bigger than themselves and their confined home environments.

Through reflecting on individuals' diary entries, Nadya built an additional understanding of what can motivate and engage people during this COVID-19 period. Specifically, the diaries gave Nadya an empathetic-understanding of, and insight into, the ways that a wide spectrum of people (importantly: most of whom were neither medical doctors nor software

specialists nor ethnographers) were reacting to the pandemic. It was undeniably clear that individuals felt *em-pathos* and yet, importantly, they also felt a deep fatigue and a sense of being overwhelmed.  Many individuals also skipped a day (or more) of the 3-weeks of diaries, and often commented on this in a later entry, writing something along the lines of "I'm sorry I didn't write anything yesterday. Honestly I was just too emotionally exhausted." At other times, when participants were asked to reflect on a particular moment of their day during which they had experienced a particular emotion like gratitude, irritation, or happiness, some individuals answered with concise phrases like "too tired. Can't think." On still other days, participants who had formerly expressed utter fatigue and emotional exhaustion, filled their pages with lengthy reflections, suggesting that their energy levels varied substantially from day to day.

Reading ~100 of such entries every night before going to bed (and often consequently resonating with the "too tired, can't think" feeling), Nadya began to see patterns of behaviour emerge. While these patterns did not revolutionise the way she currently thought about human nature, they were nonetheless striking: (1) lots of individuals experienced alternating intensities of boredom, stress, hope, fear, anxiety, and gratitude; (2) many individuals' routines had been disrupted; (3) many voluntarily expressed a desire to do something 'valuable' and 'meaningful' with their newly-acquired spare time; (4) importantly: individuals had shifting, often-unpredictable levels of energy. While it would not have felt ethical to approach these participants with an offer to use their time to volunteer on the myICUvoice project, Nadya began to realise the strong likelihood that there were potentially countless individuals sitting at home who would gladly offer a few hours of their time to contribute to something they deemed meaningful. Perhaps these individuals, like so many others had along the last several years, would indeed deem myICUvoice to be worthy of their *em-pathos* offerings.

Consequently, Nadya encouraged the five members of the myICUvoice team to reach out to their personal networks for tasks that did not require sustained periods of time commitment. While the medical doctors (Tim and Katy) and software engineers (Mary-Ann and Phil) of myICUvoice were primarily occupied with medical and software needs, Nadya began to reach out to her networks to see if there would be any significant uptake of people willing to volunteer for other tasks. She intuitively understood that the tasks would have to be something that could ideally be completed in a single session at the computer, as there was no guarantee that people's shifting energy levels would allow them to take up the same task the next day.  A summary of the groups who were reached out to in the first week (the end of March 2020) include:

- A LinkedIn Post asking specifically for graphic designers: received 25 emails of designers, animators, and illustrators offering their time
- A group message sent to the Cambridge University Women's Basketball team (of which Nadya is an alumni) : received offers to translate into Spanish, French, Italian, Polish, Turkish, and Mandarin. Team members were also responsible for recruiting native speakers who could translate into Arabic and Hindi
- A Facebook post to Hughes Hall College (of which Nadya is an alumni) : offers to translate into Portuguese
- A Facebook post to Canterbury High School Alumni (of which Nadya is an alumni) : offers to translate into German and Dutch

- A personal request to two friends: collaborated to write/edit the script, make the animated video, and do the voiceover for myICUvoice's App Release video[2]
- A personal request to three friends: to collaboratively take care of myICUvoice's social media outputs

As the weeks of various stages of local, national, and global lockdown went on, each of the core myICUvoice team members experienced their friends and personal networks responding to their candid requests for help with various tasks (mostly in the vein of language translation and/or verifying former translations.) Free legal advice was provided by Howes and Percival law firm, and, once we reached the point of gaining traction with public media, Cofinitive agreed to handle all of our press release communications. Hearing about us from our social media outputs, phonesForPatients and iComms for ICUs Project (who were both donating repurposed business iPads to ICUs in response to COVID-19) offered to pre-install myICUvoice on the iPads before they donated the iPads to ICUs.

Collectively, over 40 volunteers from several different countries had been mobilised and were crucial in bringing myICUvoice to the point that it is at now. They made animated videos, infographics, and social media posts which were then used to bring awareness to members of the public and nursing staff about myICUvoice. They also translated the app into 12+ different languages and tested it to provide feedback on how the app's content came across to a native speaker when the app was used in a non-English language. The multilingual features drew the attention of medical workers from different cultural and linguistic backgrounds, and ICU workers from several different countries including Canada, Ireland, France, India, and Australia contacted the company to request to use the app in their own care contexts. Upon being freely available on the UK app store, it was trialed by senior doctors from several different ICUs in the UK who were interested in implementing it in their own units. This interest from different medical spheres across the UK and the globe obviously did not materialise out of thin air: it corresponded to the work of 40+ volunteers from several different countries whose volunteer efforts enabled not only the software of the app to be made ready for release on the App Store, but who were also crucial in marketing the app and letting medical professionals know that myICUvoice was freely available for their use.

While, as we mentioned above, we attribute this rapid scaling to mass-*em-pathos* generated by the pandemic, we further suggest that this *em-pathos* had to, in fact, be empathetically understood. In order for the surge of *em-pathos* to become tangibly useful to the startup's scaling process, it required careful and insightful recognition, cultivation, and harnessing of it—an ethnographically-inclined vantage point and disposition was well-placed for doing this. Volunteers' *em-pathos* driven desire to volunteer had to be first and foremost 'harnessed' in a respectful manner so that no one's goodwill was abused. In this way, our volunteers became similar to the end-user of myICUvoice—not in the sense that they would end up using the finished product, but because we had to intimately understand their needs and ensure that we 'designed' a system in which they felt cared for and understood.

We did this in a number of ways. In some of the more practical ways, we broke down the tasks into relatively small tasks that could be completed within an hour or so. For example, when individuals worked on translating the app into their native language, we divided the translations into small portions so that no singular individual would feel overwhelmed. Often, we had multiple translators working on any given language, and we

used a system where they could communicate with each other (often to debate a particular translation) and with us as they translated. Understanding from her diary study that many individuals wanted to volunteer, in part, due to their own isolation, Nadya also arranged group video calls where individuals who were volunteering their time had the option of speaking live with people who were working on similar projects. During these calls, individuals sometimes exchanged ideas and advice about how they were adjusting to life at home, and there was even a time where a more senior animator and illustrator gave some free advice and training to a more junior illustrator. These conversations, often sporadically veering away from the specifics of myICUvoice to more general forms of human connection before returning to the work at hand, played a key role in ensuring that individuals felt that they were meaningfully connected to an interesting project, rather than being an anonymous cog in a wheel. There were even moments, in the midst of discussing the infographics that were being made to explain myICUvoice to nurses, that the conversations about myICUvoice shifted to offering gardening advice; and when one of the volunteer illustrators sent off a short video to Nadya to use as educational material; Nadya responded with a home-made video clip about how and where to plant snap peas. These details may seem so small that they could easily be brushed aside as meaningless, but anthropologists have been trained to deeply pay attention to the seemingly mundane and particular; it is often in the mundane that integral moments play out. Indeed, various volunteers were brought together, and sustained, not *just* because they wanted to contribute to a meaningful project, but also because they wanted to connect to others in meaningful ways during these unprecedented and often stressful times. A fine-tuned and finely trained empathetic understanding of this situation enabled the *em-pathos* of volunteers to be harnessed.

## CONCLUSIONS AND LESSONS LEARNED

Having now walked through the journey of myICUvoice in its pre-pandemic, acute-pandemic, and chronic-pandemic phases, there are some broader questions that we can return to and explore regarding the role of ethnography. We see, in the pre-pandemic timeline, during phase one and two of scaling, that ethnographic insights played an important role in understanding the nurses as a diverse set of end-users whose needs had to be understood and designed for in order for myICUvoice to gain uptake. These phases demonstrate the importance and value of involving human-centred design and ethnographic approaches early on in the process of technology development, possibly even before initial prototypes are created, for example as is done in the i-Teams programme in which Nadya and myICUvoice participated. Tim used his own empathetic understanding to identify a problem for patients, used his knowledge and experience as an ICU doctor to suggest a technically-focused solution (a communication tool), built a working prototype and showed how it improved patient experiences. As is also common in startup case studies, this prototype enabled the ICU team to improve their treatment of patients, even without that being one of their initial goals, showing the importance of putting early-stage prototypes into the hands of real end-users to assess their potential impacts. By doing so, he discovered the critical importance of nurses in accepting (or blocking) the adoption of a new tool.

It is an interesting theoretical exercise to consider whether, had a formal ethnographic approach been involved from the start asking "what is it like to be in an ICU and where are there problems to resolve?", it might have resulted in a broader early identification of the

detailed dynamics of an ICU and the key gatekeepers —though it might consequently have resulted in the identification of a completely different set of problems and a different product altogether!  Furthermore, although the original MVP was not based upon ethnographic insights, it was indeed a plausible solution to a known problem. And, by the very fact that it offered a solution to a certain set of end users, this meant that the early end-users could and did provide useful feedback based on their access to an early prototype product. Ethnographers who work alongside and with businesses must constantly negotiate and navigate these two options: (1) approaching and understanding end-users *before* a solution or product is designed and enacted and then attempting to design an ethnographically-informed solution (2) bringing forth an existing solution or product and, based on ethnographic insights and feedback, learning how to pivot an existing solution to best meet the needs.

Crucially, we note that it was the experience gained from the first version of the software, together with carefully-gathered ethnographic understanding of nurses and of the app's wider goals, that allowed the myICUvoice core team to use the pandemic-inspired increase of volunteers in a way that would drive the product forward. Afterall, it was the early version of the software which provided us with an established technical base which met the complex landscape of medical technology (including data security) before the user interface could be adapted to meet specific end-user needs. Without this background, it would have been difficult to have a suitable product which could have entered a medical context—regardless of how well it catered to the human needs.

Additionally, we speculate that, had a surplus of volunteers come before having gained an understanding of the nuances of the healthcare system and the multiple end-users and internal stakeholders, we would not have been able to successfully harness this energy into something valuable for myICUvoice's scaling vision. Indeed, it took, first, having an in-depth understanding of what the nurses needed in order to design appropriate software changes (especially the Nurse Zone) and the marketing and educational material which served to attract and secure the attention of various health professionals. As is common in many startup case studies, our experience shows the importance of being in the right place at the right time with a relevant product and a solid understanding of our potential end-users' needs and motivations.

This meant that when the environment changed in an unpredictable way (a global pandemic increasing the need for ICU beds and widespread quarantines creating a large pool of potential volunteer labour), Tim was able to respond rapidly, and his former volunteers (including Nadya) had the time to help him do so. By using Nadya's diary study to give a thorough understanding of the likely motivations and realistic levels of contributions of the potential pool of additional volunteers, the team was successful and effective in mobilising that workforce to take myICUvoice rapidly to a growth phase. Although a volunteer workforce is unlikely to be available again in such a concentrated way, early-stage startups often use an unpaid or low-paid skilled workforce due to lack of funding, and usually do so in a very ad hoc way which can lead to later loss of goodwill. The myICUvoice experience shows that the value of this workforce (and potentially its size) can be maximally leveraged by taking the time to develop an empathetic understanding of their motivations and time capacity, and project managing the required work to fit with the workforce's needs, rather than adopting the more usual approach of insisting that the workforce fit with the 'ideal' needs of the project.

Fundamentally, having had a significant increase (however short lived it was) of extra resources in the form of multiple volunteers allowed myICUvoice to get over the hurdle of having what was, in effect, a non-working version; pre-pandemic we knew what we had to change, but we did not have the resources to do it. The acute-pandemic volunteer offerings allowed us to release a first version via the App store. It even gained us the spotlight in several local and global media coverages, including being broadcast via the BBC national and World news programs as well as in the health section of their website. The continued, but reduced interest/empathic response caused by the chronic-pandemic built on the media attention to ensure continued traction with multiple early trial sites.

## Key Takeaways on Harnessing Empathy

In considering all of the particulars about the myICUvoice case study, we learned some key lessons about empathy which can be applied more broadly to human-centered design and startup initiatives.

*1. Empathy doesn't exist only in research/design/innovation teams.*
There is an urgent need for ethnographers to understand and design for how empathy engages, binds and motivates ALL participants and stakeholders in the problem-solution ecosystem that is being focused on.

*2. The conditions for empathetic-understanding, em-pathos, and mass-empathy will ebb and flow.*
Over the timeline of a project, an innovation initiative, or a longer innovation undertaking such as a startup, the levels of each three of these types of empathy that we have explored here will, no doubt, change. As ethnographers, we must become attuned to when and how different aspects of empathy can best be engaged.

*3. Don't be blind to em-pathos.*
Because so many of us in human-centered design have been carefully trained to focus on empathetic-understanding, and we thus go to great lengths to imaginatively experience and see the world from different vantage points, we must be careful to *also* train our eyes to recognise *em-pathos*. *Em-pathos* may feel like a warmer, messier cousin of the scrupulously even-handed empathetic understanding beloved by anthropologists and ethnographers: embraced by some with open arms, but (inadvertently) neglected by others. None of us cannot afford to miss out on *em-pathos*. As the boundary between 'for good' and 'for profit' becomes increasingly blurred across a range of innovation contexts—and as the burgeoning subdiscipline with anthropology, aptly called an Anthropology of the Good (Robbins 2013), continues to gain prominence—the role of *em-pathos* needs to be recognised.

And yet, as much as we are clearly advocating for the power of *em-pathos* to be recognised for what it is*,* it would be disingenuous to not also call attention to its inevitable limitations. While mass-empathy resulted in numerous volunteers whose efforts drove myICUvoice forward in many ways, the logistics and realities involved with volunteer labour simply did not provide us with the ability to continue to develop the app in a longer-term fashion. As mentioned above, many of our volunteers engaged with myICUvoice in specially designed short-term bursts; we could not have expected long-term commitment from such a range of volunteers, especially once individuals began to adjust into the pandemic routines,

and the urgent desire to volunteer their time wore off. Even the five core team members of myICUvoice have had to step back in varying levels. For example, even after volunteering full time for one and a half months, and then being supplemented for another two months by funding from ACT, Nadya recently stepped down from her role as Project Manager due to the need to secure paid employment. In other words, as we have sometimes joked amongst ourselves: *em-pathos*, despite all the aforementioned power we have attributed to it, does not pay the bills. Even with this remarkable rapid scaling that has been driven by empathy, myICUvoice still faces the challenge that all startups face: it will need to be funded and resourced in the usual way, and it will need to do this in the near future so to not lose the new momentum it has acquired.

**Nadya Pohran** (nadya.pohran@gmail.com) is an ethnographer who loves solving problems. She completed her PhD at the University of Cambridge and is passionate about contributing to work that is focused on improving society. Her current areas of focus include: (1) navigating interreligious and cross-cultural differences; (2) understanding the human side of healthtech and broader healthcare initiatives.

## NOTES

1. For further information about the types of questions that were asked to patients in the early stage of myICUvoice, refer to "Communicating with mechanically-ventilated patients: can using technology help?"
http://www.myicuvoice.com/improving-communication.html

2. To watch the full app release video, see:
https://www.youtube.com/watch?v=kHW2Osh0DiA&t=43s

3. For further information about *iTeams*, refer to www.iteamsonline.org.

## REFERENCES CITED

Alasad, Jafar A., Nazih Abu Tabar, and Muayyad M. Ahmad. 2015. "Patients' Experience of Being in Intensive Care Units." *Journal of Critical Care* 30 (4): 859.e7-859.e11.
https://doi.org/10.1016/j.jcrc.2015.03.021.

Baker T, Martin J, McCluskey R, V Navapurkar, P Polgarova, Y Shen and M Trivedi. Communicating with mechanically ventilated patients: Can using technology help? Journal of the Intensive Care Society. 2017, 17(4), Suppl:93 DOI: 10.1177/1751143717708966

Collier Jennings, Jennifer, and Rita Denny. 2020. "Where is Remote Research? Ethnographic Positioning in Shifting Spaces", *EPIC Perspectives*, May 18 2020

Denny, Rita, Susan Faulkner, Julia Haines, and Lee Ryan. "Remote Research & the Challenge of Being There." EPIC Talk 4 May 2020. Accessed at https://www.epicpeople.org/remote-research-and-the-challenge-of-being-there/

Lykkegaard, Kristina, and Charlotte Delmar. 2015. "Between Violation and Competent Care—Lived Experiences of Dependency on Care in the ICU." *International Journal of Qualitative Studies on Health and Well-Being* 10 (1): 26603. https://doi.org/10.3402/qhw.v10.26603.

Moen, Ellen Klavestad, and Dagfinn Nåden. 2015. "Intensive Care Patients' Perceptions of How Their Dignity Is Maintained: A Phenomenological Study." *Intensive and Critical Care Nursing* 31 (5): 285–93. https://doi.org/10.1016/j.iccn.2015.03.003.

Pulman-Jones, Simon and Amy Weatherup. 2019. From 'Cool Science' to Changing the World: The Opportunity to Support Pre-startup Science Commercialization through Ethnography and Human-centered Design. *2019 EPIC Proceedings* pp 211–236

Robbins, Joel. 2013. "Beyond the Suffering Subject: Toward an Anthropology of the Good." *Journal of the Royal Anthropological Institute*. Volume 19, Issue 3. pp. 447-462. https://doi.org/10.1111/1467-9655.12044

Samuelson, Karin A.M. 2011. "Unpleasant and Pleasant Memories of Intensive Care in Adult Mechanically Ventilated Patients—Findings from 250 Interviews." *Intensive and Critical Care Nursing* 27 (2): 76–84. https://doi.org/10.1016/j.iccn.2011.01.003.

Tambiah, Stanley. Culture, 1985. Thought, and Social Action: An Anthropological Perspective. Harvard University Press.

Topçu, Serpil, Şule Ecevit Alpar, Bilgi Gülseven, and Ayda Kebapçı. 2017. "Patient Experiences in Intensive Care Units: A Systematic Review." *Patient Experience Journal* 4 (3): 115–27.

Wade D, Als N, Bell V, et al. Providing psychological support to people in intensive care: development and feasibility study of a nurse-led intervention to prevent acute stress and long-term morbidity. BMJ Open. 2018;8(7):e021083. Published 2018 Jul 23. doi:10.1136/bmjopen-2017-021083

World Health Organization. 2020. "Timeline of WHO's Response to COVID-19." 29 June 2020 Statement. Accessed on 14 July 2020. https://www.who.int/news-room/detail/29-06-2020-covidtimeline

# Empathy, More or Less

## Scaling Intermediary Experiences of Emotion and Affect in Innovation

LORA V. KOYCHEVA, *Technical University of Munich*

*Abstract: Questions of scale permeate current approaches to empathy in applied human-centered work—and especially design thinking—but they have remained largely unquestioned. What is more, empathy has become an empty signifier, and empathizing is often a near-formulaic and pro-forma endeavor. To catalyze a reworking of the concept, in this paper I synthesize what has been said so far of empathy and its role in design and innovation, and I take stock of what these contributions point to. I ask: "How can we think of empathy as a scalar phenomenon and thus re-scale it in innovation?" I offer some illustrative, if unresolved, tensions with empathy I have had in my own ethnographic work with a robotics start-up, and I conclude the article with a series of provocations with the hope they will be taken up further.*

## INTRODUCTION

Empathy has quickly become one of the most familiar buzz words in the world of business, whether in product design or human resources, customer research or in cross-functional team building. To invoke empathy in the contemporary professional setting is to signal "human-centered" and cutting edge. In recent years, since it was launched into the design and innovation professional vocabulary and then expanded to gain a foothold in both management and entrepreneurship milieus, empathy appears to have become not only a celebrated and desired tool into the innovation toolbox—almost a dispositif in a Foucauldian sense—but also so prominent that its invocation and application seems to have become mandatory. What was once a fresh reminder to business people, designers, and engineers that feelings, perspectives, and emotions, and not only numbers, have an immediate value to their operations and work, today appears to be fast becoming an empty signifier.

Our community is certainly not oblivious to this. As empathy´s celebratory potential began taking on formulaic and mandatory overtones, recent debates around it have offered an increasing array of counterarguments against the use of the term. In this catalyst paper, I do not aim to provide a single concluding vote to either of these camps, but rather to explore to what extent and with what effects could we, as a community of scholars-practitioners, rework the role of empathy in ethnographic research in applied settings. I synthesize what has been said so far of empathy and its role in design and innovation, and I take stock of what these contributions point to. More importantly, in so doing I am looking for clues as to whether there are paradoxes or unresolved tension in the ways empathy has

been conceptualized and deployed in our practice which might provide a fresh analytical ground for asking new questions about empathy and how it is used in applied research.

What follows is a preliminary rewriting of the question of empathy. in the first, conceptual part of this paper, Synthesis, I ask: "How can we re-scale empathy in innovation?" As a first step in this endeavor, I suggest that to empathize is a scalar activity—a point whose implications and potentials are largely lost in both practice and writings on the topic.

My position here is animated by my own grappling with the topic, some theoretical, some stemming out of my 24-plus months of fieldwork with a radical innovation moonshot venture pursuing the development of humanoid robots, where I first started out researching questions of identity and team culture, and then became increasingly involved, through participant observation, into development and outreach questions of how to create empathy for robots at societal level. In Part II, Exegesis, I illustrate the limits of empathy in innovation and the study thereof as an example to such grappling.

However, such an understanding of the power and prominence of the affective dimensions of empathy requires that we understand empathy not on a flat scale, as a temporary adoption of a worldview perspective from a point-to-point individual-to-individual (as in a "the researcher" empathizing with "the user"). It requires, rather, a more granular understanding of empathy on a nested scale, one implicating historical, cultural, and social aspects in active interplay with each other, and empathy´s reconceptualization as an inhabiting of affective states and in terms of intermediary experience of the multiplicity of its constitutive affective variants (such as, among others, hope, anger, pain, passion, fear, exhaustion, bravery, weirdness, friction). Ultimately, it allows us to better capture, conceptualize, manipulate and responsibly account for questions of scaling feelings and perspectives in our work. In Part III, Catalysis, I suggest a non-exhaustive list of provocations that might help us reframe the question of empathy.

## PART I: SYNTHESIS

Empathy´s meteoric rise to prominence in and dominance of the vocabulary and mindset of the world of design and innovation is part and parcel of the changes design thinking brought in the 1990s (e.g. Leonard and Rayport 1997). As one of the first and most distinct steps in design thinking—back then a novel approach on how to identify and solve problems—the rise of empathy as a concept and as a fundamental step in the innovation process in the last 20 years can easily be pointed to as one of the true success stories of a long-standing and continuously ongoing push for peopling engineering practice and management thinking. As a result, recent decades have seen a substantial number of professionals adopting it as their occupational identity and becoming empathy coaches, empathic strategists or empathy gurus, and entire dedicated "empathy labs" exist both as independent businesses and within large corporations such as Google and Facebook (Stinson 2020). In taking stock of the merits of empathy as part of design and innovation, as well as the challenges and dangers posed by its increasingly near-automatic and formulaic application lately, we must tack back and forth between not only what the term means and what it does, but also place it within a larger understanding and increasing critique of design thinking as the leading framework for innovation.

Originally starting as merely a new product development framework, but then steadily expanding into questions of customer centricity and organizational culture, design thinking practitioners lay claim to have a clear map of "applying the principles of design to how people work" (Kolko 2015; see also Kolko 2014; Brown 2009). The claim was that it created better outcomes, more finely attuned to user needs and "pain points" (Platzer 2018) than hitherto delivered by a remote bird´s-eye view of quantitative approaches. Part of its revolution has been to bring decision-making and product testing outside of the confines of labs and into the real world, placing the designer not only as a creator of specifications and aesthetics deduced from their own imaginary about the world in which their creations will be embedded and will circulate, but also as a validator and generator of real-life insights on how such a potential product would be experienced, and—crucially—understanding and placing the perspective of the user and the user´s reality above one´s own assumptions.

A key differentiator that design thinking claimed for itself in its approach was a kind of empathic perspective-taking that other approaches lacked, catapulting empathy—meaning "in feeling" from the Greek pathos via the German Einfühlung—as the go-to method of tapping into other people´s realities via sharing their inner experiences. What "being in feeling" meant produced a number of definitions, sometimes full of contradictions, which I am about to suggest, points to the weaknesses of adopting empathy as an approach – weaknesses which we should be either collectively moving away from, in favor of more ethnographic thinking, or working to eliminate and make stronger.

Thus, Battarbee et al. have defined empathy as "the ability to be aware of, understanding of, and sensitive to another person´s feelings and thoughts without having had the same experiences" (2014, 2 my emphasis), while a little later in the same text, they suggest and affirm practical approaches to achieving empathy precisely through experience-near techniques, such as, for example, to "participate in grueling endurance events to share athletes´ exhilaration and pain" (2014, 4), recalling to mind Lois Wacquant´s call for embodied methods, an "incarnate study of incarnation by practical example" (2014, 4).

Renowned product designer Jon Kolko describes it thus:

> "empathy is about acquiring feelings. The goal is to feel what it´s like to be another person. That goal is kind of strange, because it´s unachievable. To feel what someone else feels, you would needs to actually become that person. You can approximate her feelings, so product research intended to built empathy is really trying to feel what other people feel. Assuming you aren´t actually an eighty-five year old woman, consider for a second what it feels like to be an eighty-five year old woman. This consideration is still analytical, it´s about understanding. You need to get closer to experiencing the same emotions that an eighty-five year old woman experiences, so you need to put yourself into the types of situations she encounters [to] approximate her feelings, leaving your own perspective in order to temporarily take on hers" (Kolko 2014,5)

Michael Ventura similarly notes, "empathy is about understanding. Empathy lets us see the world from other points of view and helps us form insights that can lead to new and better ways of thinking, being, and doing" (2018).

In sum, if one were to approach the concept of empathy as championed by design thinking (e.g. Brown 2009) and applied in marketing and leadership contexts (e.g. Ventura 2018) and product development ( e.g. Kolko 2014), the promises that approaching the lived

realities of those for whom we design, share moments of (cross) "cultural intimacy" (Herzfeld 2005), and for whom and with whom we ultimately create value are so many, that surely we should subscribe to deploying empathy without second thought. As described in these widely admired and popular approaches, empathy promises straightforward and surefire ways into other people´s realities and offers the ability to quickly and amenably tap into exactly the tranches that we need to understand for the purposes of delivering insight. Those applying empathy are implicitly portrayed as swiftly deploying it as a tool—although, unlike in ethnography, we never see this—a tool which works magically to translate the immanent and immaterial (feelings and lifewords) into the profitable, the material, and the immediate—objects, structures, services. As Jennifer Wong quips cheerfully (or ironically?) in an online article on creating empathic design systems, "...to help solve the UX process problem, inject a bit of empathy" (2019).

Whether it is presented as a tool from the designer´s toolbox (Kolko), as a mindset and a way of being (Ventura), or even perhaps as a medium to be administered (Wong), the one aspect which all proponents and interlocutors of empathy and design thinking agree on is depth: the prize is to understand "deeply" (e.g. Stinton 2020; Kolko 2014) and to achieve "perspective", often seen as the product of "stepping in other people´s shoes."

For anthropology, on the other hand, the question of accessing, understanding, and representing, in a formulaic shorthand, "how they feel in their shoes," has never been a simple affair. The discipline has dealt with the question of fellow feeling as a vehicle to knowledge and as a subject of inquiry in a characteristically discerning manner. It has examined the question of "fellow feeling" (Solomon 1995), and has recognized a difference between empathy, emotion, and affect as three distinct domains, all of which require various levels of engagement with context, focus on embodied experience, and in which narrative and language mediate what is essentially an intersubjective experience that is both slipping, and yet firmly enmeshed, within social and political imperatives and structures (see, for example Lutz and White 1986; Besnier 1990; Beatty 2013; Beatty 2014 for comprehensive reviews; and on affect, Skoggard and Waterson 2015; Stodulka et al 2018; Newel et.al 2018). What is more, the ambiguities and limits of knowing "other people´s minds" has been shown to be always linguistically mediated (e.g. Keane 2008), but also necessarily embodied.

Thus both Daniel White (2017) and Danylin Rutherford (2016) have suggested that affect is largely unspoken and involves an embodied intensity of feeling which in turn gives rise to emotion within the subject. White succinctly captures the historical shift in the field between emotion and affect: "if anthropologists of emotion throughout the 1970s and 1980s had shown how feelings variously fix and stick through different compositions of language and discourse, anthropologists of affect shortly thereafter sought to show how some feelings slip, evade, and overflow capture" (2017, 175). In other words, if empathy is the ability to bridge inter-personal varieties of existence in the search for capturing meaning, it requires a reorienting of cognitive, affective, and bodily states.

Clifford Geertz´s famous skepticism as to whether adopting "the" native´s point of view is analytically valuable comes to mind here, as he argues instead for a "hopping back and forth between the whole perceived through the parts" (1983, 69). This is a subtly scalar proposition of engaging phenomena on a nested scale, and not a singular point-to-point one. Numerous other scholars have further unpacked the density of the concept. Famously, Renato Rosaldo´s poignant essay "Grief and a Headhunter´s Rage" (1993), on understanding murderous grief after the loss of a loved one only after the tragic death of his wife during

fieldwork, suggests that there are domains of human experiences which are viscerally comprehensible only to those who have gone through them. More recently, and in a different vein, Douglas Hollan (2008)has argued that empathizing is an intersubjective act not only of feeling but also of imagination—and, crucially, is not the work only of the one empathizing but also requires a reciprocity of emotion and imagination on the part of the one being empathized with. This last point suggests that empathy is a perspective-taking exercise based not only on a singular agent, but is rather the product of two agents taking perspective with respect to each other—meeting on a mutually re-scaled perspectival plain. Finally, C. Jason Throop suggests that "empathy…must always be understood in the context of particular cultural meanings, beliefs, practices, and values… it is significant to explore how empathy is both recognized and enacted by individuals in its marked and unmarked forms but also to examine the specific contexts, times, and situations in which empathy is possible and valued and those in which it is not" (Throop 2010, 772; also Hollan and Throop 2008).

Yet "standing in their shoes" and "seeing like they are seeing" has been deemed increasingly deceptively formulaic.EPIC community members have already put forth a range of thoughtful objections to the preeminence of empathy discourse. Rachel Robinson and Penny Allen (2018), for example, have argued compellingly that empathy is not to be conflated with evidence, and have discussed the many traps in which they perceive empathy can introduce unwelcome and unhelpful bias . Tamura and colleagues (2015) have demonstrated that a "sense of ownership" is much more effective in the innovation and entrepreneurship context than empathy in that it creates more powerful research . John Payne (2016) has commented on Paul Bloom's (2017) recent arguments against 'empathy' as a decision-making rationale. Payne carefully examines the limitations of empathy, noting: "Many of these methods have been repurposed from the social sciences to the needs of design practice. However, when removed from their theoretical foundations and optimized toward identification of *user needs*, they don't account for the social implications of the work we do. This needs to change" (2016). Romain, Johnson, and Griffin (2014) have been similarly preoccupied with the ways in which empathy obscures the potentially meaningful to consider tensions between stakeholders in business. Finally, in an even more provocative vein, Thomas Wendt (2017) has argued that empathy is too human-centric, reductive in its Western anthropocentrism, thus essentially rendering the political aspects and questions of power in design essentially invisible, to the detriment of all.

In sum, for professional ethnographers, the way empathy is approached in most design thinking is problematic, stemming from an increasing tension between design thinking and ethnography. As Jay Hasbrouck has elegantly pointed out, design thinking has become "symbiotic in practice, but […] at odds empirically" (2018, 3) with ethnographic approaches, creating an unwelcome conflation between the kinds of questions that design thinking can ask and answer, and those that ethnographic thinking can, in addition to inaccurately framing all human-centric approaches as reductive.

But lest we consider that it is the anthropologists who are particularly critical of the concept of empathy, skepticism of it and its application has also been mounting in parallel in design circles. Without mincing words, Natascha Jen has spoken against it, calling design thinking as a whole "B.S." for being too prescriptive, and signaling out empathy as specifically problematic: "the word empathy is prevalent in design discourse; people have become experts on design empathy. Back in the day we called it research…" (2018). And, in

what is perhaps most damning condemnation because it comes from one of the most authoritative voices in design, Don Norman (2019) has admitted to not believing in empathic design for several reasons. One is because of the inherent inability to design for "many" by immersing yourself in the individual experiences of the one or the very few; another, because very often people´s own understandings of their own experiences and feelings are not immediately accessible to themselves. Finally, in his view, the ways empathy and human-centric design operate at present, they simply cannot solve for the truly wicked problems, such as climate change and hunger, for example—something that Natasha Iskander critiquesd  in the pages of the Harvard Business Review as the inherent tendency in design thinking to protect the status quo and to reinforce the position of the designer, but not designed for—even if empathy is employed, because 'solving for' is the remit of the powerful (2018).

What emerges as a pattern, then, is that although empathy remains a fruitful, popular, and profitable approach to obtain perspectives and mine them for understandings of experience, its shortcomings are increasingly being exposed. Key among them are that it does not address its political potential and is regularly ahistorical; it does not lend itself readily to understanding contexts defined by uncertainty and complexity; it ignores key questions of the positioning of subjects—including in relation to each other; it can get lost in translation between research encounter and the production of an object. It does not make a critical distinction between reported experience and shared experience; and fails to explain how it deals with the limits of verbal explanation. Further, it does not differentiate critically between cognitive and affective empathy in a systematic manner, or explain when to use which variant. Nor does it address well how empathy operates from within the fraught entanglements of objective and subjective phenomena. Pain is one such phenomena, ironically enough. David Platzer (2018) has given the concept of "pain point" an excellent treatment. However, when the question of what the pain point means is refracted through a careful consideration of the role of empathy in it, it becomes necessary to situate both at multiple scales and levels of analysis: one objective (there is in many cases such a thing as real experience of physical pain, discomfort or unease which innovation addresses) and the subjective, more elusive forms of experiencing them—something which C. Jason Throop, not incidentally also thinking about pain, has termed "intermediary forms of experience " — "much of what we deem to be experience is characterized by … transitions, margins, fringes, by the barely graspable and yet still palpable transitive parts of the stream of consciousness that serve as the connective tissue between more clearly" (2009, 536).

In sum, although an inherently relational phenomenon, both in its reliance on accessing other people´s experiences and in translating them into different metaphysical forms (be they objects and services that circulate often locally and globally), current approaches to empathy fail to factor in something which anthropologists have long understood, examined, and theorized: emotions and affect are as social as they are cultural, and they are socially constructed, always enmeshed at the nested scales of individual and society, always rife with political potential, and always refracted through questions of meaning and power, always contextual, fleeting, incomplete, and elusive.

A good way forward, I believe, is to take a cue from anthropology's  insistence on unpacking what perspective is, and to think of perspectives precisely as scalar phenomena, and, in turn, scales as being a question of perspective and positioning. This is a running theme in both branches of approaches to empathy: the anthropological one and the design

Empathy, More or Less – Koycheva

thinking one. In many ways, then, both anthropology and design rely first and foremost on taking perspective, which is an inherently scalar phenomenon, as a recent edited volume on the topic has proposed. Drawing on Marilyn Strathern´s definition of scale as "the organization of perspectives on objects of knowledge and enquiry" (2004, xiv in Summerson Carr and Lempert 2016, 5), E. Summerson Carr and Michael Lempert argue in the introduction to their edited volume on the pragmatics of scale that "scaling involves vantage points and the positioning of actors with respect to such vantage points means that there are no ideologically neutral scales [and that] scaling is process before it is product" (2016, 3-4).

Noting various examples of scale from a range of cognate disciplines, from distinctions between "private," "personal" and "political" to "macro" and "micro," and even conceptualizations such as "bench-to-bedside" throughout their introduction and the volume, a key motivation is to show "the inherently perspectival nature of scale, asking of our material "whose scale is it," "what does this scale allow one to see and know" and "what does it achieve and for whom" (2016, 15, original emphasis). In a subsequent chapter, Susan Gal further highlights the comparative logic inherent in both scale and perspective: "scaling implies positioning and, hence, point of view: a perspective from which scales (modes of comparison) are constructed and from which aspects of the world are evaluated with respect to them" (Gal 2016, 91).

Yet in borrowing from anthropology, and re-scaling the process of perspective taking for business contexts in making it faster, less granular and less concerned with language, context, and embodiment as constitutive of empathy, design thinking has lost the kind of granularity that is exactly what makes empathizing a very rare kind of empirical tool for understanding other people´s realities.

Coupled, however, with the proliferation of the discourse of empathy in the business milieu, it would appear that two camps are forming. One is calling for more empathy—scaling it qualitatively and championing a more granular and extended research at the empathy step in the innovation process—and the other is signaling that the concept has ceased to be analytically useful. Where does that leave our field?

I propose that instead of seeking to substitute one´s own perspective for that of the user in attempting to gain perspective through a "like" state, a more ethnographically informed approach to gaining perspective would be to pursue a "with" state. Instead of "seeing like them," "seeing with them" allows practitioners to position the otherwise wicked problem of capturing and representing others´ experiences in a granular manner by situating the empathizing endeavor at multiple scales at ones. In the next section, I offer two illustrations on the challenges for so doing, and in the final, catalyst section, I briefly touch upon what the opportunities might be if the field takes a turn in this direction.

## PART II: EXEGESIS

Admittedly, the grapples that inform my provocations on the need to re-work the concept of empathy into more granular variants are borne out of work different to the commercial projects that are often presented here at EPIC, which focus on issues such as UX, product development, and organizational culture. Rather, the context of my (originally purely academic) work is extreme (cf. Hallgren, Rouleau, and de Rond 2018) in that it is unique and cannot be said to represent most commercial settings in which applied ethnography operates. Specifically, I work with a moonshot startup which dwells uneasily in

the space between the commercial demands and expectations of the venturing scene and the scientific requirements and realities of research: an academic venture occupying the outer extreme edge of an already extreme category of innovation, which Sarasvathy (2008, 93) has termed "the suicide quadrant"—where a new product is introduced into a new market. In the case of humanoid robots, it is largely the case that the innovation is so radical, that there is no product, no urgent demand, and no immediate market. Traditionally, this has meant that either only large companies such as Google can afford to have in-house units (such as X, the Moonshot Factory) dealing with such kind of innovation, or that the government gets involved (cf. Mazzucatto 2011). In the case of my collaborators, neither of these were not the case—thus making product development and keeping the startup financially afloat a Herculean task. It faces the kind of slow diffusion and challenging scaling based not so much and exclusively on kind of innovation which does not rely on the quick diffusion cycles of lean driven product development but requiring the slow but steady interpretative understanding of how to disrupt meaning as a necessary early ingredient (Haines 2016).

Yet it is precisely this far-off vantage point that gives me a different vista on questions of how we approach empathy, affect, emotion, and experience more broadly in the search for useful understanding of others. This approach draws on the strengths and contributions of academic anthropology´s unpacking of these questions to which I referred in the previous section. But it also transforms questions from being meaningful and relevant into also being applicable and interventional. Applied ethnography makes such a pivot in its daily operations, which nonetheless do not preclude the ability to draw on and contribute to theory equally well. This point is worth insisting on given that we are still collectively working to end the "theory-practice apartheid" (Baba 2005).

I never intended to study questions of radical innovation, let alone musculoskeletal humanoid robots. Rather, as a scholar I was interested in how startup teams form in the academic context, and how their identities and practices inform team culture. But as I was studying questions of culture, identity, and practice within the setting of a moonshot startup in the academic setting , as is often the case with prolonged fieldwork, I became more and more incorporated into the team, slowly and over the course of many months, through our shared understanding that a sociocultural anthropologist has a legitimate role in a startup developing humanoid robots, especially where sociocultural outreach is concerned.

To be sure, no single paper could capture the multiplicity of angles through which the topic of empathy as a scalar and perspectival project, rather than as simply a method to gain perspective, is refracted in every milieu conceivable in innovation and entrepreneurship. In what follows, I offer two vignettes from my own ongoing work, which serve here to illustrate why I am compelled to question empathy in design thinking. The first instance revolves around questions of the robot´s features and appearance, and questions of gender and race in particular. The second vignette draws on how an unexpected failure of empathy resulted in developing one of the most popular pre-programmed function the robot has: hugging. In both instances, I chart the dilemmas that empathy, as a scalar, perspectival, embodied, and linguistic phenomenon, presents to our current thinking on the topic.

## Cute, White, and Boyish? About a Roboy

Robots are regularly evaluated on their utility: what they can do. This is true for industrial robots but also for humanoids, robots that attempt the visage and shape of human

beings. The humanoid robot Roboy (fig.1) was conceived as a different kind of humanoid robot, not only because of the technological intricacy of developing the corpus of the body with a musculoskeletal mimetic engineering solution, but also notably because of the vision driving the robot´s development—a robot whose body is as good as a human's—and its initial raison détre, a positive messenger of artificial intelligence.

Roboy is not only a functional mechatronic system, however. With an extensive dialogue system, the robot can enter in conversation and display a number of emotions, such as smiling, winking, blushing, frowning and sadness (fig.2).
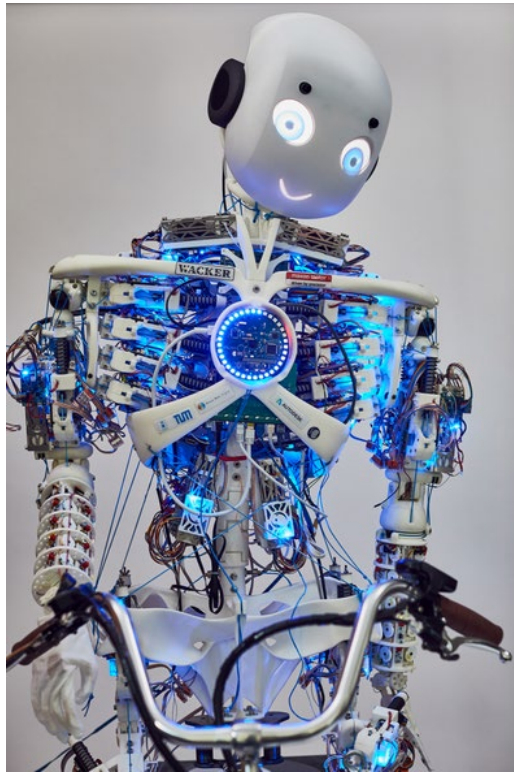


Fig. 1 Roboy 2.0 smiling interface © Devanthro –The Roboy Company. Used with permission.

My arrival at Roboy coincided with the unveiling of the second generation robot (fig. 1 and 2), Roboy 2.0, affectionately referred to as the "big brother" of the original robot developed at the University of Zurich, Roboy Junior.

Through a number of participant observations at international fairs and national and local events, which I attended largely to observe how such a unique team works and comes together at professional events, I nonetheless managed to document how the public interacts with both Junior (now forever retired) and 2.0. A key difference between them was that Junior was small, immoveable, and non-interactive, whereas 2.0 was a towering robot with a range of pre-programmed interactive facial expressions (fig. 2).
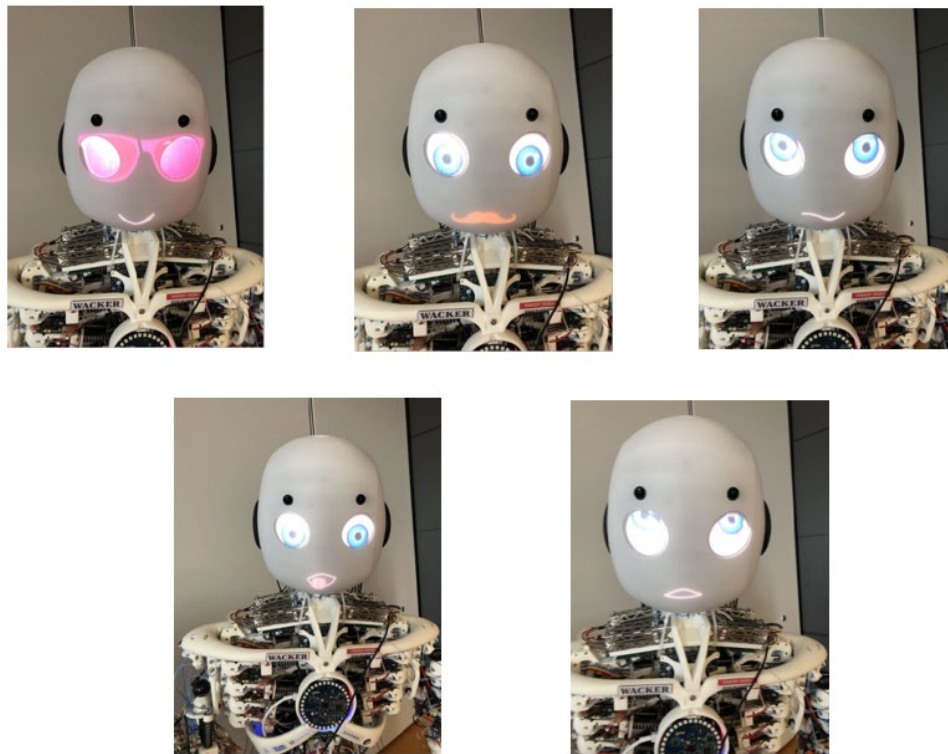
Fig. 2 Range of expressions on Roboy 2.0. Compilation. © Devanthro –The Roboy Company. Used with permission.

Again and again, people coming in contact with the robot made comments like "What can it do?," "Why is it a boy?" and "How cute." Only rarely, but importantly, did people make timid remarks about the robot such as "oh, it´s white."

Important things in thinking about empathy in relation to the robot include how it scales in understanding how people react to a humanoid robot, how they would react if a particular feature was changed, and also how to make the robot empathetic in his interaction with people. The first one is not too surprising, given that the original meaning of the word 'robot' is rooted in the word for slave and that our collective understandings of robots continue to run along the lines of utility, as noted above. The question of the representation of gender and the emotional responses people have when interacting with the robot, however, as well as the remarks about color, are crucial here, because they bespeak a perspective according to which the robot is seen and interacted with. This is a perspective that the development team and the CEO must take into account, but which is not always easily integrated with the vision for the robot.

To say that people have to empathize with the robot so that it is accepted as a positive messenger for robotics and AI is too facile and belies the fact that people will empathize with it from their own particular perspectives and embodied experiences. For example, the question about the robot's gender is overwhelmingly asked by women—a sign not only of

Empathy, More or Less – Koycheva

the times we live in, in which female empowerment and representation is as under attack—but also of female members of the public working out a way to identify with the robot. Remarkably, the fact that while the robot is a "boy" and appears somewhat boyish his voice is female is almost always lost on the public.

At some events we have done together, the CEO somewhat peevishly explained that the robot was developed initially by "9 dudes in a lab at the University of Zurich" and that creating a fun near life size female figure—an over-sized doll, as it were—would have been not only weird but also inappropriate.

Similarly exasperating but fascinating are people's comments on the color of the robot. Technically and literally, the robot is "off white", a result of the standard material used in 3D printing. But the CEO has been asked many times to 'correct' this by making it whiter, and he has resisted. But while this off-white color could also be a wonderfully playful way of reaffirming the old semiotic axiom that because we are all different, we are all the same, this is not what the public wants or has seen. Rather, the question suggests that people tend to identify on a flat, literal, one-to-one scale, and that they are approaching the visage of the robot not only from a perspective of identifying with it, but also from within a larger contemporary culture that encourages personalizing everything. And the robot, in demanding that it has its own techno-selfhood, resists personalization.

Thus the robot presents a grand design thinking challenge in which empathy becomes fundamental in its perspectival and scalar dimensions, and where the question of empathy becomes anthropological in that it requires that you draw on knowledge of culture, context, and discourse, in addition to trying to step in the proverbial shoes. How do you build a humanoid robot with human features and make him a symbol of positive robotics if the identification and representational issues are not possible to integrate in a single design? If you want your robot to be human—in the universal sense of humanity—how do you make it particular enough to satisfy the immediate need for people to identify with it along gender and racial lines? Should the robot be personalizeable? How would that change how we collectively think about robots?

The question of the scales of the universal and the particular confound and escape the narrow logic of empathy as a state of changing perspectives with which we have come to grapple here. Urgent updates on the term are needed.

And in addition to that puzzle, there is the question of how cuteness scales. Cuteness is currently one of the most valuable propositions for the robot, seeing that the technology is so difficult to develop that there are no immediate markets yet. It is also a cornerstone concept around which one of the revenue streams of the robot revolves: the team regularly exhibits new technology being implemented in Roboy for sponsoring partners at international fairs.

One of the key challenges currently facing the team is how to scale the appearance of the robot so that it appeals to the widest possible public, not only in terms of race and gender, but also, given that its human-like expressive appearance is one of its key defining features, one which differentiates it from all other humanoid robots being developed. Cuteness is a property that children possess, but as the robot is advancing in generation and its range of manipulative abilities are further developed (currently Roboy can grasp but not walk yet), the question of its appeal moves more towards the center as pivotal to tackling questions of how acceptable musculoskeletal humanoid robotics will become in society in future. Scheduled to be unveiled some time in fall 2020, Roboy 3.0 will have morphed into

yet another version of a humanoid robot. As the robot grows in functionality, should he age? Will the boy forever remain a boy? As a socio- cultural interface, one that is built to promote empathy between robots and humans and which relies on human empathy, these questions are as imminent for the team as they are wicked. They are questions that demand anthropological intervention in scaling meaning, above all: questions that require empathy at multiple scales, and going beyond singular use case scenarios and singular user understanding.

## Hug by a Robot

The text comes at lunch time, saying that things at the international fair where the Roboy team is exhibiting with a partner, are going dismally. The engineers are underslept, disheartened, and "bored", the sponsors are "concerned" and "unhappy," and the public is passive and disengaged. This came on the heels of a really rough, pressure cooker week of development which overshot its very hard deadline, where emotion was high in the team, tempers flared up and stakes were mounting to deliver on a technical challenge that suddenly was not working out in the last moment, despite prior successful tests. That the fair was going down the drain is the last thing both the CEO and I wanted to hear: him as the leader and key responsible person, standing at the front of the booth with his corporate partners; me as the ethnographer who has been immersed in the team for some time now, working alongside people whose dedication and passion for their work I have come to admire and draw inspiration from myself. I know already that the team will emerge forever changed from the last week. Turbulence is not over, and this is unwelcome and worrisome news. I ask what the problem is.

Stuck in the midst of a tricky dance that anyone who has had to plan and execute an international fair exhibition knows: how to both fit in and differentiate your product at the same time with little budget, a partner signaling being underwhelmed, and a team who were working through the night to deliver against the odds, the CEO was facing an ostensible wall: "Basically, too many robots now. So just a Roboy is no longer enough."

Thousands of miles away, back at my desk, I close my eyes and I try to imagine how it feels at these fairs that I have observed before, and what the team is going through right now. What works and what doesn't. In my mind´s eye, I see the endless stream of only half interested people, whose glances you are trying to catch; the uncomfortable bumps of bodies around booths that have managed to gather attention, and the true awkwardness of sitting there with nobody in front of your booth. The scantily clad and heavily made up women selling robots at competitors´ booths and the giant culinary extravaganzas with free food and drinks. I can almost hear the myriad conversations surrounding a person at all times, sloshing into one giant wave of low murmuring sound. I can almost feel the stale air. In my mind´s eye, I can see the team, each and every one of them in the mood I have come to know well at fairs. One of them is so pure an engineer, and not at all a salesman, that he refused to see the value of going to these "boring as f***" events, where even the free food at other booths could not be tempting enough for him to justify why he is forced to stay there and not be back in his lab, "doing the real work, the work that matters": developing. I see another member of the team, focused and razor sharp, doing what must be done, because in her words, this is her "life", and this is her "family." Always composed, she makes sure that the robot and its subsystems function as expected either directly or by

delegation, but also keeps an eye out to ensure that the rest of the engineers are completing the necessary tasks. I see two of the more junior team members: one who is willingly suffering the many fast micro interactions with the steady stream of people with whom he is discussing the robot, and one who is responsible for stress within the team, having underestimated the amount of prep he needed to deliver a fundamental piece of tech—something for which he has tried to make up for by sleeping only two hours a day, and working the rest of the time. And I see Rafael, the CEO, continuously interfacing on behalf of his own company and on behalf of his sponsors, whose latest tech his robot is there to showcase. He is equally under pressure, from all sides, and trying to solve the problems typical of the remit of the CEO, which none of the other team members would face. I try to imagine how he feels and how they feel. Logically I can understand that he is bored, worried, and underslept, but at my air conditioned desk, having slept just fine while they were probably still programming and soldering at 3am in prep for the next day, I lack the visceral experience to relate.

I focus on how I first felt when I saw the robot at a fair and I tried to adopt the perspective of a bored engineer milling around from booth to booth, all fairly indistinguishable from one another. I remember I was intrigued but intimidated to touch it. It had wires exposed in the plastic exposed rib cage, perching usually on a bicycle or a pedestal, looking wobbly and ready to break at the smallest touch (something which rarely happens, and which the Roboy engineers routinely address by inviting people to physically interact with the robot, a prompt which does not work most times).

I try to think what unites these very different people, and of all the various feelings that this less-than-a-minute interaction foments. There is the curiosity of what it is, and why is the robot "so big." The hesitation to ask, for fear of sounding stupid. The hot stale air pressing in a person, the unfresh bodies surrounding you. The odd discrepancy by the sheer tangle of obvious cables (perhaps a more daunting sight for non-engineers than engineers) and the wide blue-eyed winking face of the robot.

Tech. Engineering. Curiosity. The tinkering spirit. That is what unites them. So I type back after a while:

> "organize a play session : )". "use social media -  #engineersatplay #fairsarenotboring #robotsarecool #flash #make #hack – to invite people who are bored as fuck by flyers and chicks to actually tinker with something. 1. Assuming tech crowd are full of engineers who are bored by the static and slow nature of everything and are [there] because [the location is fancy] and 2. [the partner] gets traffic. Social media gets clicks and likes, Roboy gets to surprise and stay agile. Nobody expects a guerilla hackathon at a buttoned up event."
> "hmmm," comes back a text. "#robotsareboring I like, but then what´s the complement…#becomearobot? Bearobot. indefinitelife. Playforever. Exploreforever. So the play idea is excellent. I was already going in that direction"

A while later, another text comes, in the disorienting fashion of sudden text, when a person is plunged into the *medias res* of another person´s realities.

> (Rafael)  "they´re stoked. So good timing."
> (Lora) "who?" I ask. "How?"
> (R) "team"

(L) "roboy?"
(R )"to try extreme measures. I´m bored."
(L) ": )"
(R) "so I want to have fun"
(L) "play time"
(R) "yes"
(L) "guerilla hackathon.nobody expects"
(R )"yes.not as easy"
(L )"ohhh…keep me posted"

Later on that day, it turns out that Rafael had decided to stage a hug-a-robot demo instead, and the team was busy with coding in the necessary procedures. I check in on the next day in the afternoon, curious to hear how things were going at the booth: "hows it working? The strategy"

(R ) "amazing. like no other"
(L) "really? Interesting. Send proof. I have very serious doubts."
(R) "interesting"
(L) "you know why right?"
(R) "no"
(L) "because of the great threshold which usually exists, and that has been observed at fairs, on how to interact with Roboy…normally, people are terrified of precisely touch…unlike with Junior, there they usually touch either the hand or the face, because the robot is smaller and more childlike"
(R ) "it´s even crazier. Roboy sits on a pedestal"
(L )"maybe you standing there, beautiful handwriting sign and winning smile in place, directing people to hug the robot works"
(R) "no roboy holds the sign of course"
(L) "so do they actually come and hug?"
(R )"yes"
(L) "interesting"
"who?"
(R) "all sorts of people"
(L) "women or men more? Age range? What do they say? Do they linger before they approach?"

Soon thereafter, I find myself added to a specially programmed Telegram messenger channel, in which the robot is documenting the number of times he has hugged a human in real time (fig.3) "u even wanted proof" comes the half accusatory, half triumphant text by Rafael.
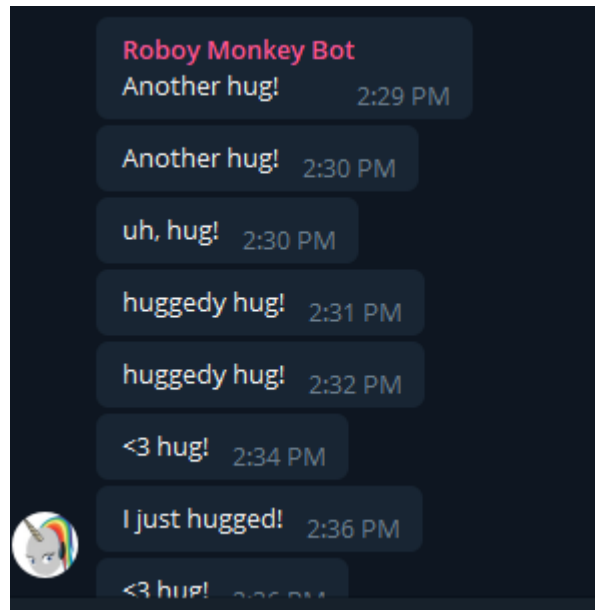
Fig. 3. Screenshot of real-time hug capture channel. Own capture from screen.

"of course," I text back. "I am a scientist. about this of all things. *hugging* a
*robot* and a public anonymous space. I want all the proof in the world. Because it
will be very important—what do people say?"

Hugging, as it turned out, was so successful, that it became a cornerstone of every
exhibition ever since, and a key marketing value proposition for booking the robot at fairs.

It is important to note here that had the CEO made a decision based on my empathy
with the bored engineers in the audience (organize a hackathon), he would have gone in a
misdirection, and in many ways, he made a decision rooted not in empathy – and fairly
despite it. Although he reported that the audience, his team, and himself were bored, and
although I gave a recommendation solving precisely for that, ultimately the decision to settle
on hugs was drawn from larger sociocultural preoccupations on how to make robots social.

His decision, based not out of empathy with this team, but of a need to create empathy
for the robot, managed to achieve both – in creating empathy for the robot in offering free
robotic hugs, he also managed to rally his team, deliver value for his sponsors, and work
towards culturally acceptable human-robot interactions. Although intuitively executed, in this
case empathizing was placed within a nested scale of multiple converging relations – human-
robot, public-team, team-manager, team-sponsors.

## PART III: CATALYSIS

Traditionally, papers end. Conclusions rephrase what has been said until now, and tie up
any loose ends. But this being a catalyst paper, the task here is to use the final stretch to ask
myself, in the company of your patient attention this far, what it is that we should be
catalyzing in the world of innovation, applied ethnography, humanoid robotics and AI, and
scales; that is, the scales of disciplines and the scales of empathy. I have synthesized literature

from both sides of the empathizing endeavor (anthropology and design thinking), and I have offered an illustration on how empathy has appeared in my work in the realm of robotics innovation. I have suggested that empathy is a scalar endeavor and that scales are perspectival in turn: to take a perspective is to re-scale, but that hardly ever involves a singular plain of action, but rather operates on nested logics. In lieu of concluding, however, instead of tidying this up, I would like to pick out and leave hanging loose ends for our community.

I do this via a series of intellectual provocations. These provocations are rooted not only in the synthesis and interpretation of the literature, both applied and theoretical, which I have offered. They draw not only on my musings of what empathy is, what it could be, what its limits are, and how can we rescale it, but also on the inevitable limits of my datasets and my experiences. These are questions that I cannot immediately take up, but I hope that we all collectively will.

One question to consider might be: what is lost and what is gained if the inquirer is always "on" emotionally in the pursuit of an empathic understanding? Do Malinowskian moments of un-grace, when we are impatient with our respondents, when we are at odds, or simply when we get them wrong or disagree profoundly with them illuminate situations and reveal additional dimensions of any particular situation which we are trying to understand? What extra dimensions of understanding does this add in the context of innovating from w human-centered perspective?.

What are the limits of empathy and can going against it have a positive outcome? In a broad vein, I have called for a move away from an attempt to achieve "like" states, or—at the very least—a healthy suspicion towards them. Yet, instead, could we conceive of empathy as a phenomenon dependent on spatio-temporal adjacencies – what I have called "with" states— those not of switched and temporarily replaced perspectives but as a space of motion oscillating comparatively between perspectives – a space of parallax as a process (Ballesteros 2015) as it were, not of perspectives as stable states, heeding Geertz´s call to "tack back and forth"?

Could we conceive of empathy as a space of negotiation and translation—embodied, linguistic, political and ethical—and how would that improve the ways in which we deliver actionable results?

What is lost in the lack of affective and evocative writing whenever the ethnographic account is replaced by the executive summary? In other words, when the ethnographic methods of taking perspective taking are decoupled from the painstaking explanation of the ethnographic account of how this came about in any particular human-centered encounter, how can empathy be accounted for? Alternatively, what new forms of ethnography, ones derived not from traditional fieldworks but from collaborative practices in a varieties of practical settings, can emerge in future?

How do we "take the perspective" of non-human actors or entities that nonetheless demand innovation´s attention: phenomena such as climate change or urban renewal, or robots and self-driving vehicles, as well as in context of interspecies relations?

Another question to consider as we move into more and more digital contexts is how we empathize in the digital realm. How is the lack of fully embodied experiences, otherwise so necessary to empathizing and gaining perspective providing a challenge for empathy in innovation, especially as we are, as of the time of this writing, continuously besieged by a

raging pandemic which has forced an increasingly virtual online social interaction and life upon the world?

What is lost and what is gained in allowing for empathy to become a trend, and almost an ideology of innovation? How can we salvage from the trivializing hype the role of relating in our work?

What is the relationship between empathy and time? In the rapid contexts of business, how could empathy ever be achieved?

Finally, in going forward, how can design address the relationship of empathy and power, in acknowledging its interventional potential (Suchman 2011)? How will that help us deliver better insights to our customers, and help us create socially responsible businesses for the 21st century? Instead of attempting to see and feel and experience like, say, a black woman within an interview or even a day in order to create a product for her, one can, however, hire or collaborate with one—and in the space between these two perspectives, not only novelty but also ethics can be born.

I have suggested here that empathy is a scalar phenomenon, but that this aspect of it is inherently lost. The time is right to rework empathy not as a facile "standing in other people´s shoes" but as a negotiated, complex phenomenon of relating and of taking up positions, which are as political as they are experiential.

My final and most important point is that perhaps there is no need for designers to rescale the complexity of feeling for their purposes, and in the process lose precisely what makes ethnographic insight so powerful and valuable: its granularity and its ability to relate in embodied, situated, contextual ways.

Rather, designers should collaborate more with ethnographers—so we can have our empathy, and scale it too.

**Lora V. Koycheva** is a sociocultural anthropologist working on questions of innovation, entrepreneurship, and robotics in society. You can best contact her via Linkedin: https://www.linkedin.com/in/lora-koycheva/


## NOTES

## REFERENCES CITED

Ahmed, Sara. 2014. *Cultural politics of emotion*. Edinburgh, UK: Edinburgh University Press.

Baba, Marietta. 2005. "To The End Of Theory-Practice 'Apartheid': Encountering The World." In *Ethnographic Praxis in Industry Conference Proceedings* vol. 2005 (1): 205-217. https://doi.org/10.1111/j.1559-8918.2005.tb00023.x

Battarbee, Katja, Jane Fulton Suri, and Suzanne Gibbs Howard. 2014. "Empathy on the edge: scaling and sustaining a human-centered approach in the evolving practice of design." *IDEO.* Accessed September 8, 2020. *http://www. ideo. com/images/uploads/news/pdfs/Empathy_on_the_Edge. pdf ..*

Beatty, Andrew. 2013. "Current emotion research in anthropology: Reporting the field". *Emotion Review.* 5(4), 414-422. https://doi.org/10.1177/1754073913490045

------------. 2014. "Anthropology and emotion". *Journal of the Royal Anthropological Institute*, 20(3), 545-563. https://doi.org/10.1111/1467-9655.12114

Besnier, Niko. 1990. "Language and affect." *Annual review of anthropology,* vol. 19(1): 419-451. https://doi.org/10.1146/annurev.an.19.100190.002223

Bloom, Paul. 2017. *Against empathy: The case for rational compassion.* Random House.

Brown, Tim. 2009. *Change by Design: How Design Thinking Transforms Organizations and Inspires Innovation.*[Kindle 2 version]." *Retrieved from Amazon. com .*

Gal, Susan. 2016. "Scale-making: Comparison and Perspective as Ideological Projects." In *Scale,* edited by M. Lempert& E. Summerson Carr. *University of California Press, Berkeley,*91-111.

Geertz, Clifford. 1983. *Local knowledge: Further essays in interpretive anthropology.* Basic books.

Haines, Julia Katherine. 2016. "Meaningful Innovation: Ethnographic Potential in the Startup and Venture Capital Spheres." *Ethnographic Praxis in Industry Conference Proceedings.* Vol. 2016 (1) https://doi.org/10.1111/1559-8918.2016.01084

Hasbrouck, Jay. 2017. *Ethnographic thinking: From method to mindset.* Routledge.

Hällgren, Markus, Linda Rouleau, and Mark De Rond. "A matter of life or death: How extreme context research matters for management and organization studies." *Academy of Management Annals* vol. 12 (1): 111-153.

Herzfeld, Michael. 2005. *Cultural intimacy: Social poetics in the nation-state.* Psychology Press.

Hollan, Douglas, and C. Jason Throop. 2008. "Whatever happened to empathy?: Introduction." *Ethos* 36 (4): 385-401. https://doi.org/10.1111/j.1548-1352.2008.00023.x

Hollan, Douglas. 2008. "Being there: On the imaginative aspects of understanding others and being understood." *Ethos* 36 (4): 475-489. https://doi.org/10.1111/j.1548-1352.2008.00028.x

Iskander, Natasha. 2018. "Design thinking is fundamentally conservative and preserves the Status Quo." *Harvard Business Review* 5 .

Jen, Natasha. 2018. "Design Thinking is B.S." https://www.fastcompany.com/90166804/design-thinking-is-b-s Last accessed September 27, 2020.

Keane, Webb. 2008. "Others, other minds, and others' theories of other minds: An afterword on the psychology and politics of opacity claims." *Anthropological Quarterly* 81(2): 473-482.

Kolko, Jon. 2015. "Design thinking comes of age." *Harvard Business Review* 66-71.

-------------.2014. *Well-designed: How to use empathy to create products people love*. Harvard Business Press.

Leonard, Dorothy, and Jeffrey F. Rayport. 1997. "Spark innovation through empathic design." *Harvard Business Review* 75: 102-115.

Lempert, Michael., & E. Summerson Carr, 2016. "Scale: Discourse and dimensions of social life" In *Scale. ,* edited by M. Lempert& E. Summerson Carr. *University of California Press, Berkeley,*

Lindholm, Charles. 2005. An anthropology of emotion. A companion to psychological anthropology: Modernity and psychocultural change, 30-47.

Lyon, M. L. 1995. "Missing emotion: The limitations of cultural constructionism in the study of Emotion". *Cultural Anthropology* 10(2), 244-263.

Lutz, Catherine., & Geoffrey White, 1986. "The anthropology of emotions." *Annual Review of Anthropology*, 15(1), 405-436.

Mazzucato, Mariana. 2011.*The Entrepreneurial State: Debunking Public vs Private Sector Myths.*

Norman, Don: Why I Don´t Believe in Empathic Design. https://xd.adobe.com/ideas/perspectives/leadership-insights/why-i-dont-believe-in-empathic-design-don-norman/

Payne, John. 2016. "What's so Funny 'Bout Peace, Love, and (Empathic) Understanding." Retrieved September 10, 2018.

Platzer, David. 2018. "Regarding the pain of users: towards a genealogy of the "pain point"." In *Ethnographic Praxis in Industry Conference Proceedings*, vol. 2018 (1)301-315. https://doi.org/10.1111/1559-8918.2018.01209

Roberston, Rachel and Penny Allen. 2018. "Empathy Is not Evidence: 4 Traps of Commodified Empathy." In *Ethnographic Praxis in Industry Conference Proceedings*, vol. 2018 (1) 104-124. https://doi.org/10.1111/1559-8918.2018.01199

ROMAIN, T., JOHNSON, T. P., & GRIFFIN, M. (2014, October). On Empathy, and Not Feeling It. In Ethnographic Praxis in Industry Conference Proceedings (Vol. 2014, No. 1, pp. 181-182).

Rosaldo, Renato. 1993. "Introduction: Grief and a Headhunter´s Rage." In *Culture & truth: the remaking of social analysis: with a new introduction.* Beacon Press.

Rutherford, Danilyn. 2016. "Affect theory and the empirical." *Annual Review of Anthropology* 45: 285-300. https://doi.org/10.1146/annurev-anthro-102215-095843

Sarasvathy, Saras D. 2009. *Effectuation: Elements of entrepreneurial expertise.* Edward Elgar Publishing.

Skoggard, Ian., & Alisse Waterston 2015. "Introduction: toward an anthropology of affect and evocative ethnography". In *Anthropology of Consciousness*, 26(2), 109-120.

Solomon, Robert C. 1995. *A passion for justice: Emotions and the origins of the social contract.* Rowman & Littlefield.

Stinson, Liz .2020."The Empathy economy is Booming, and There´s More at Stake than Just Our Feelings, Part 1 and Part 2." Accessed September 27, 2020. https://xd.adobe.com/ideas/perspectives/social-impact/empathy-in-design-economy-good-bad-ugly/.

Stodulka, T., Selim, N., & Mattes, D. .2018.. Affective scholarship: Doing anthropology with epistemic affects. Ethos, 46(4), 519-536.

Suchman, Lucy. 2011. "Anthropological relocations and the limits of design." *Annual Review of Anthropology* 40: 1-18. https://doi.org/10.1146/annurev.anthro.041608.105640

Tamura, Hiroshi, Fumiko Ichikawa, and Yuki Uchida. 2015. "Goodbye Empathy, Hello Ownership: How Ethnography Really Functions in the Making of Entrepreneurs." In *Ethnographic Praxis in Industry Conference Proceedings*, vol. 2015 (1) 322-330. https://doi.org/10.1111/1559-8918.2015.01058

Throop, C. Jason 2009. "Intermediary varieties of experience". In *Ethnos*, 74(4), 535-558. https://doi.org/10.1080/00141840903202116

---------------------. 2010."Latitudes of loss: On the vicissitudes of empathy." *American Ethnologist* 37 (4): 771-782. https://doi.org/10.1111/j.1548-1425.2010.01284.x

Ventura, Michael. 2019. *Applied empathy: The new language of leadership.* Hachette UK, 2019.

Wacquant, Loïc. 2014. "Homines in extremis: What fighting scholars teach us about habitus." In *Body & Society* 20(2): 3-17.

Wendt, Thomas. 2017. "Empathy as Faux Ethics." Retrieved January 4, 2019 from https://www.epicpeople.org/empathy-faux-ethics

White, Daniel. 2017. "Affect: An Introduction". *Cultural Anthropology* 32 (2):175-80. https://doi.org/10.14506/ca32.2.01

Wong, Jennifer. 2019.. The How, and Why, of Creating an Empathetic Design System. Accessed September 27, 2020.  https://xd.adobe.com/ideas/principles/design-systems/creating-empathetic-design-systems/ .

# Dissent at Scale

Dissent assumes a common ground, a common purpose. How do we push back against scaled-up discrimination and disinformation when they come from different cultural logics?

Session Curators: Nick Agafonoff, Lindsay Ferris, Tabitha Steager, Erin Taylor

PECHAKUCHA

# I'm Not Scaling, the World Is Really Scaling Against Me

## Or, What Will 4 People Think / Chaar Log Kya Kahenge

SMRITI KAUL, *Convo Research & Strategy Pvt Ltd*

*This paper raises the implications of simplifying algorithms for scale and uplifting content that is damaging for human evolution. Technology is powerful because of its scale and also disempowering for the same reason. Scale is in the variables and online media, in the zest of empowering women, is deciding our fate. I get it when the housewife looks to YouTube to cook a meal. However, I also see the heartbreak when what should be freeing is actually being used to throttle progress. When a girl from a small sub-segment of global population like Rajasthan, while wanting to feel empowered realises that she's unable to measure up? Are we responsible for this? Are our "hashtags" and "likes" fuelling our continued repression?*

*As an ethnographer, I study media consumption to overcome barriers to participation in the online world, and as a gender trainer, I also create and use media content to overcome barriers in the real world. I find myself continually curious about what we learn and how we replicate. Hence I'm concerned that the models we have for scale aren't healthy especially for girls and women. Can we consider a shift from quantity to quality of scale - from how many to how? Can we find a path that broadens the evaluation for scale? What might be the indicators for scale that progress society?*

4 people that are now 4 billion - Is scale broadening what matters or dictating how I think?

**Smriti Kaul** has a Masters Degree in Law with a specialisation in Human Rights. Her passion for the subject led her to social and behaviour change communication research, content development, impact evaluation and gender training. She has an experience of 10 years in development sector with Government of India, UN, SAARC , Girl Rising, Action Aid, among others. As an ethnographer with Convo, she loves doing studies which brings deep insight into mindset and behaviour of young teenagers reflecting interesting dynamics of their socio-cultural environment. She also enjoys doing studies that bring out impact of gender on individual choices and preferences.

CASE STUDY

# Fighting Conspiracy Theories Online at Scale

REBEKAH PARK, *Gemic*
DAVID ZAX, *ReD Associates*
BETH GOLDBERG, *Jigsaw*

*This 2019 project conducted in the US and the UK sought to understand which conspiracy theories are harmful and which are benign, with an eye towards finding ways to combat disinformation and extremism. This case study demonstrates how ethnographic methods led to insights on what "triggered" conspiracy belief, the social and emotional roles conspiracy theories played in believers' lives, and how conspiracy belief was often a reflection of a person's sense of societal alienation. We discovered that any extreme interpretation of a conspiracy theory could be harmful. The findings of this project changed how the client — and by extension the developers behind major tech platforms — understood harmful conspiracy-related content. The aim of this project was to inform how to scale and amplify the work of individual conspiracy debunkers.*

**Keywords:** *Conspiracy theories, fieldwork, engineers*

## INTRODUCTION

In 2019, Jigsaw, a technology incubator within Google, and ReD Associates, a strategy consultancy, undertook ethnographic research on conspiracy theorists across the United States and the United Kingdom. The project set out with the initial mandate to understand which conspiracy theories are harmful and which are benign, with an eye towards finding ways to combat disinformation and extremism online. Although a small cadre of self-motivated conspiracy theory "debunkers" generate content online, their efforts are insufficient to tackle the proliferation of conspiracy misinformation online—some of which motivates serious violence (The August 2019 El Paso shooting of 23 people in a Walmart was fueled in part by a belief in the "white genocide" conspiracy theory).

In its constant aim to navigate between the tension of undue restraints on speech and harmful speech, Jigsaw (and more broadly, Google) stood to benefit from being able to surgically parse harmful conspiracy content from the harmless; that way, only the harmful could be penalized. More generally, in Google's quest to better understand niches within a user base of over two billion, an ethnography of conspiracy theorists stood to render rich portraits of dimly understood and often reflexively vilified Internet users to those who broadly shape some of the Internet's most popular services.

This case study demonstrates how ethnographic methods led to insights on what "triggered" conspiracy belief, the social and emotional roles conspiracy theories played in believers' lives, and how conspiracy belief was often a reflection of a person's general sense of societal alienation.

Our initial assumption - that some conspiracy theories were more harmful than others because they could excite acts of violence - was ultimately revised, for two reasons. First, we found that any conspiracy theory, if followed to an extreme length, could become harmful.

Second, we found that the more useful outcome from our ethnography was not that there were *types of theories*, but rather *types of theorists*. At a certain point in our study, we pivoted and sought to help identify the types of conspiracy theorists that are more likely to respond to at-scale technological deterrence strategies.

By focusing on two specific ethnographic encounters, we demonstrate why it is more important to distinguish between types of theorists rather than types of conspiracy theories. Our conclusion is that "extreme" theorists themselves cannot be affected by debunking content, because they will not consider factual argumentation at all. Extreme theorists are distinguished by their visceral, emotionally driven beliefs. Rather, debunking content is best deployed to people who are milder in their conspiracy belief, at a stage where the belief has not yet become embedded and visceral. In-person ethnography was essential in arriving at this understanding, since the personas and behaviors revealed by our in-person visits often overturned the personas and behaviors suggested by a conspiracist's digital presence.

The findings of this project changed how the client — and by extension engineers behind major tech platforms — understood harmful conspiracy-related content and how to scale efforts to curtail extremism fueled by conspiracy theories.

In this paper, we begin by providing background on debunking as a strategy to dissuade people from upholding conspiracy theories and explaining why the methodology was based on an ethnographic approach. Second, we share two cases from our fieldwork to illustrate our main argument that the most strategically feasible way of combatting conspiracy theories requires us to segment different types of theorists. Lastly, we discuss how our findings impacted the way Jigsaw approached users who consume conspiracy theories online. This case study stands as an example for why ethnographic research on what happens offline helps explain, contradict, and influence what happens online. This perspective is necessary to developing and designing for online products and understanding the users themselves.

## BACKGROUND

Belief in conspiracy theories, particularly in the U.S., is not new, but the internet has made it possible to spread fringe beliefs rapidly, widely, and efficiently (Merlan 2019). Conspiracy theories continue to spread online at an alarming rate. Believers in extreme versions of conspiracy theories are sometimes moved to action. For instance, in 2016, a gunman stormed a pizza parlor in Washington DC, convinced— because of conspiracy theories circulating online— that it was the site of a child sex trafficking ring. Understanding the line between a playful conspiracy theory and one that motivates people to harmful action is crucial. So, too, is understanding what can be done to help debunk conspiracy theories in a way that is persuasive, so that people who start to fall down conspiracy rabbit holes can climb back out.

Jigsaw has studied misinformation for years, but conspiracy theories caught their attention as a less well understood form of misinformation that was closely, and repeatedly, linked to real world violence. They began focusing on the questions of how conspiracy theories could be so powerful that they motivate violent action, and how to deter conspiracies. A popular approach used to counter conspiracy theories is debunking. This typically entails engaging others one-on-one in great depth, or sometimes via broadcast, to counter very specific, and often highly technical, arguments. Jigsaw looked for efforts to

scale debunking and came across the work of Mick West, an expert at conspiracy theory debunking.

Mick is a successful video game programmer (noted for his role in the popular Tony Hawk skateboarding series) who retired early and became a full-time debunker. He's the author of a book titled *Escaping the Rabbit Hole: How to Debunk Conspiracy Theories Using Facts, Logic, and Respect / A Guide to Helping Friends, Family and Loved Ones.* We also consulted the foundational scholarly literature on conspiracy theories, including *The Paranoid Style of American Politics* by Richard Hofstadter, and *Conspiracy Theories* by Cass Sunstein and Adrian Vermeule, whose notion of the conspiracy theorists' "crippled epistemology" we employed in our analysis.

It was Mick's book that was the touchstone, though; in it, he outlines a process that involves taking seriously the points offered by the believer and offering counterinformation. Rather than being dismissive, he brings a deep sense of empathy, a wealth of knowledge, and tremendous amounts of patience and care to each interaction.

Mick has been debunking conspiracy theories for years, and he's an inspiration for his deeply empathic approach. He has made dozens of videos on YouTube and runs a forum called Metabunk where he hosts debates on theories as varied as 9/11, Chemtrails, and the notion that the moon landing was a hoax. His book contains a few great success stories of people who have been *deep* down their rabbit hole, but have gradually been coaxed out.

Still, Mick West is only one man. And even though there are other conspiracy debunkers online, too, the problem is simply too big for a handful of hobbyist debunkers to make a real dent in.

A large portion of society believes in a conspiracy theory to some degree - in the 2015 article, "Conspiracy Theories and the Paranoid Styles of American Politics," political scientists Eric Oliver and Thomas Wood found that in any given year, about *half of the American public* endorses at least one highly dubious conspiracy theory. No matter how popular Mick West's websites become, this painstakingly personalized approach to debunking simply can't scale to meet demand for this challenge.

This raises the question, how does one "scale" debunking? And is it even possible, or achievable in ways whose costs do not outweigh its benefits?

## STUDYING CONSPIRACY THEORISTS

Academic research on conspiracy theories, limited to the Western, English-speaking context, has largely focused on the psychology of individuals who believe in conspiracy theories and why they believe in them (Kluger 2017, Preston 2019, Roose 2019, van Prooijen and van Vugt. 2018). Psychological factors exploring why certain individuals are motivated to uphold conspiracy theories highlight universal traits that make one receptive to this type of content, or how the belief in conspiracy theories is reflective of other existing psychological behaviors. For instance, a variety of cognitive differences were found to increase susceptibility to conspiratorial thinking, such as schizotypy, paranoia, or delusional ideation (Dagnall, Drinkwater, Parker, Denovan, and Parton 2015). Individuals with cognitive differences are engaged in a world where conspiracy theories have explanatory power. Believing in conspiracy theories can fulfill emotional goals by feeling good about the world or exerting a feeling of control and order amid feelings of powerlessness (Hart 2018; Kluger 2017; Imhoff and Lamberty 2016; Grzesiak-Feldman 2013; Whitson and Galinsky

2008). Also, social exclusion may lead people to support conspiratorial beliefs because they provide social meaning and value (Graeupner and Coman 2017). Psychologists have argued that people who were from low-status groups (less education and wealth) were more likely to believe in conspiracy theories (Douglas et al 2019; Freeman and Bentall 2017). What we found particularly relevant to this study is how one conspiracy theory acts as a gateway to other conspiracy theories. Once a person accepts one conspiracy theory, he/she is more likely to be receptive to other conspiracy theories (Brotherton, French and Pickering 2013; Jolley and Douglas 2014; van Proojien and Douglas 2018).

Beyond psychology, scholars of media studies have examined the role that social media has played in spreading conspiracy theories and helping to form new types of communities online (Jolley and Douglas 2014; Stempel, Hargrove, and Stempel III 2007; van Prooijen and Jostmann 2013). Despite the fact that conspiracy theories found online are theoretically accessible to anyone, researchers have found that conspiracy theory content tends to stay within specific communities that are already receptive to it or are actively seeking conspiracy theory content (Douglas et al 2019). One study found that conspiracy theories, in the case around the Zika outbreak, was not spread online through a central authority but rather through a series of decentralized networks (Wood 2018). This suggests that people share and consider conspiracy theories outside of, or separate from, "official" stamps of approval or authority figures.

Our research builds upon existing studies in three ways. First, we focused on the role context, or large-scale social, economic, and political factors, play in shaping conspiratorial worldviews. This is distinct from psychological perspectives that are primarily focused on types of cognitive profiles that make one susceptible to conspiracy theories. We probed further into how the environments that people were living in were connected to the formation of conspiracy worldviews. For instance, conspiracy theories positing that because a small group of elites control the global economy helped people understand their lack of social mobility. Second, similar to the research on the instrumental nature of conspiracy theories, our research is focused on the generative nature of adopting a conspiratorial worldview. Beyond fulfilling the need for control and power, we explored the social aspects of engaging in a conspiratorial worldview, including making friends, having a sense of purpose, and feeling excitement when theorizing with others. Third, we investigated the relationship between what people do online vs. offline. We traveled to meet theorists in person to gain a wider view of their everyday lives in their homes and workplace. We sought to understand how people go from consuming conspiracy theory content to acting upon it. We define "acting upon it" as everything from forwarding a website to others, to liking a post, to meeting other believers in the local library, to openly considering, on theoretical grounds, harming the culprits responsible for conspiracies.

The goal of this study was to learn whether a relationship exists between types of conspiracy theories and potential for harm. Even though a correlation between extremism and belief in conspiracy theories exists, it is not the case that believing in conspiracy theories will automatically lead to extreme or violent behavior (Bartlett and Miller 2010). We conceptualize harm broadly to include the individual consequences on their personal relationships (e.g. estranged parents), health (e.g. refusing cancer treatment), and social status (e.g. being outed as a white nationalist). We also define harm in terms of actions taken against others who believers blame as perpetuating or benefiting from conspiracies, such as

immigrants. In addition, we also considered but did not directly investigate the harms that conspiracy theory belief has upon faith in governments and institutions (Coaston 2018).

**METHODOLOGY**

To better understand how to stem the tide of false and potentially harmful conspiracy theories, Jigsaw had to better understand how people came to hold a conspiratorial worldview, what conspiracy thinking does for them and their lives, and what people do, if anything at all, with their beliefs. We wanted to understand how conspiracy theories fit into their overall life and what role conspiracy theories played in motivating other actions, offline. While it is an important area of study to understand the psychological factors that explain how a person even comes to believe in a conspiracy theory, we were more focused on tracing life histories and identifying any patterns between belief and action, specifically with an eye toward harms that are linked to believing in conspiracies. Rather than investigating levels of education and intelligence or cognitive deficits, we wanted to understand contextual, circumstantial, and personal factors that led someone down the rabbit hole, as well as what factors kept them from falling deeper. What role did the people around them, or life events, play in upholding or backing away from theories?

To answer these questions, our team of five researchers conducted in-person, in-depth interviews with 42 conspiracy theorists across the US and UK, as well as expert interviews with academics and journalists investigating conspiracy theories. In accordance with our initial hypothesis that some conspiracies were harmful and others innocuous, we recruited respondents across three different conspiracies: two theories we believed could be tied to real world harm and a third "control group" theory we believed was likely to be harmless. In the "believed harmful" camp were theorists who believed in "false flag" events (the notion that, for instance, mass shootings have been staged—which has been linked to harassment), as well as believers in "white genocide" (the notion that immigration trends indicate a deliberate plot to eliminate whites—which has been linked to mass shootings). In our "believed harmless" camp were believers in various science-related conspiracies (e.g. chemtrails, flat earth).

We used websites like 4chan and Twitter as starting points for recruitment and observations. By searching for the term "white genocide," for instance, or a related term called "the Kalergi Plan," we were able to follow, converse with, and ultimately recruit participants over Twitter. We also used surveys with questions designed to screen for conspiracy belief, and we drew from our personal networks as well. Since many conspiracy theorists are of course skeptical of strangers, going through intermediaries was often helpful.

Prior to conducting research, the team familiarized themselves with media coverage and scholarship on conspiracy theories and interviewing experts. We actively cultivated having an open mind and took particular care in understanding what language to use, and how to present ourselves in ways that would not cause us to lose our credibility (e.g. showing our familiarity with theories online rather than boasting about academic credentials). We consciously did not use the phrase "conspiracy theories" because of the negative connotations, but rather spoke about "alternative narratives", "research," and "truth." Our approach was to be honest in presenting ourselves as former academics and journalists and offered our sincere interest in understanding and listening to their points of view, and to

have them guide us through the websites, videos, and channels that they used as sources of information. We also offered to meet in public places and did not record or take pictures, unless permission was granted.

It should be noted that in the end, all of the people we met with were not any different than the types of people we usually meet in other types of studies. Our research participants represented a variety of fields including teaching, technology, construction, and healthcare. Building rapport with them was similar to any other interactions we have in the field, though with more awareness around language, being actively empathetic, and not drawing suspicion with recording devices.

Ultimately, we made in-person visits, each lasting several hours, in or around people's homes. (One of our researchers also explored the lighter side of conspiracy culture by attending the "Storm Area 51" event held in October.) The insights gathered through fieldwork were analyzed alongside literature on grappling with the history of, or sociological studies of, conspiracy theories, including: *Kill All Normies*, *Fantasyland*, and *Republic of Lies: American Conspiracy Theorists and Their Surprising Rise to Power.*

## OUR FINDINGS: DISTINGUISH BETWEEN THE THEORISTS, NOT THE THEORIES

As discussed, our initial hypothesis was that certain *theories* are more extreme or harmful than others. In other words, that a person's likelihood to commit harm was related to the type of conspiracy they believed. We had an assumption that pseudo-scientific conspiracy theories like flat earth or chemtrails—the belief that the government is spraying mind-controlling chemicals from planes—were relatively innocuous. Meanwhile we assumed that racially-tinged theories like white genocide were perhaps dangerous by definition.

But what we found surprised us. We learned that it was less important to distinguish between *theories*, and more important to distinguish between *theorists*. What matters is how much of a person's life is taken over by a conspiratorial worldview. If everything is part of the conspiracy, a person can no longer trust anything or anyone. An extreme conspiratorial worldview frames the elite "they" as powerful and as the enemy. Thus, it is not surprising that some studies have found that belief in conspiracy theories could be a predictor of having committed crimes or stating that they would commit a crime (McNamara 2019; Herrema 2019). In our own study, believers of extreme versions of conspiracy theories justified killing the enemy, if they knew who it was, because they would be saving others. It was not a particular theory that drove people to action but rather how much that person lived within an extreme conspiratorial worldview.

*All* conspiracy theories, we came to learn, have the potential to be harmful — more on that in a moment. And when it came to conspiracy theorists, we found, there was a very wide spectrum in terms of how hardened a person's conspiracy belief is. This becomes very relevant to know when you are hoping to try to "debunk at scale."

To explain why, we will first discuss a trip the three authors made to Montana.

## DEEP DOWN THE RABBIT HOLE: THE HARDENED THEORIST

In November 2019, we flew to a remote town in Montana to meet some friends of friends who believe the earth is flat. We met a couple who bonded over conspiracy theories

and attend weekly meetups where they discuss and "test" conspiracy theories (for instance, by pointing telescopes at the horizon to try to determine the shape of the earth).

The woman we met with, whom we will call "Jennifer," grew up with hippie parents who moved the family deeply off the grid. Homeschooled through youth, by her mid-30s Jennifer was living with her parents in a remote corner of Montana, without internet reception. It was only through her gig house-sitting that she was able to access the internet at all — which was how she met "Carl," her first romantic partner, on an environmentally-themed dating website.

Soon, Carl began sending Jennifer thumb drives full of conspiracy material that she could consume on her home computer. Jennifer mainlined hours and hours of videos; down Carl's rabbit hole she went, and by the time we met her, her idiosyncratic beliefs were legion. She believed the Earth was flat and there was a nefarious agenda to mislead us about its true shape. She believed that Hitler was a great man and that the Holocaust didn't happen. She believed the government or the cabal controlling it sprayed mind-weakening chemicals from airplanes. There was hardly a conspiracy theory we had encountered that Jennifer *didn't* believe in. Her whole worldview had been reprogrammed, doubly so now that Carl had moved to Montana to be with her. She, Carl, and others in their remote area began hosting a weekly conspiracy meetup, that doubled as something of a self-help group.

What we came to sense, meeting with Jennifer and people like her, was that there was no such thing as an innocuous conspiracy per se. *Any* theory could become dangerous or extreme, depending on what other theories it was caught up in. Jennifer's flat earth belief was intimately tied up with the idea that a cabal of people—likely Jews, she had come to feel—were lying to her about the nature of the world. Given the right opportunity, she said, she would attack a representative of this cabal — in fact, she said, she would consider such an act a form of "self-defense" due to the "scale of the atrocity" this cabal was perpetrating on humankind. No longer did flat earth belief appear to us to be inherently innocuous. Belief in a flat earth, as discussed earlier, is a reflection of how extreme a conspiratorial worldview is. To believe in a flat earth, one must discount whole fields of expertise, such as physics and geography. In addition, to pull off a flat-earth conspiracy, one must also consider all of the instruments and narratives that propagate the theory, such as books taught in school, professors of astronomy or physics or biology.

We learned something else from meeting hardened theorists like Jennifer: that for people who have reached this hardened stage of conspiracy belief, debunking isn't the right strategy at all. The notion of debunking presumes a sort of rational, civil debate, where each side shares facts in a sporting fashion, and some victor emerges. But for people like Jennifer, the very notion of what was a "fact" had become subverted. Any mainstream source of information was now reflexively dismissed as lies; if *The New York Times* (and its happens-to-be-Jewish ownership) toed the party line on the earth being round, could it really be trusted?

Furthermore, coming away from our meeting with Jennifer, we felt that a fundamentally *cognitive-and-rational* intervention like debunking was likely to fail against someone for whom conspiracy belief appeared to serve an *emotional* role. The conspiracy belief helped them make emotional sense of a world where they felt marginalized, disenfranchised, and alone.

The intervention needed to pull such a person away from conspiracy belief is likely multi-facted, not to mention more intimate, personal, and personalized than social media can offer. Jennifer may well pull herself out from her rabbit hole someday, but it will likely take a perfect storm of personal and even societal factors before she is ready to do so.

The question of whether debunking can be scaled is moot with the set of conspiracy theorists like Jennifer: even if it could be, it won't work on her.

## AT THE TOP OF THE RABBIT HOLE: THE BUDDING CONSPIRACY THEORIST

This isn't to say, though, that debunking has no purpose. Because we did also encounter success stories in the field, showing that with the right interventions, tech platforms can help *prevent* the spread of misinformation at scale. The key, we came to feel, was to make sure debunking was deployed at the right—early—moment in a budding conspiracist's journey.

For an example of that, let's talk about the case of Lois.

When we met her, Lois, who lives outside of San Diego, believed in so-called "chemtrails." When airplanes fly at high altitude, the exhaust from airplanes cause condensation in the air — these familiar streaks in the sky are called contrails.

But proponents of the chemtrails conspiracy theory believe that in many cases, the lines in the sky aren't just water condensation, but are some sort of nefarious chemical, likely sprayed by the government. In some variants of the theory, it's all an experiment in climate control. In other variants, the chemicals are poisons that conspirators use to subtly undermine the will of the population, along with fluoride in the water.

When our researcher met Lois at an Italian restaurant in a San Diego strip mall, she explained why she believed in chemtrails. "I've seen them!" she said. She explained that back in 2015, her brother, a rancher, had pointed them out to her. Her brother said the government must be spraying poisons to "control the masses." That struck Lois as a little far-fetched, but she went home and started doing internet searches related to chemtrails. She went to NASA's website, but couldn't find anything debunking it. Instead, she eventually landed a video with a number of pilots and other self-proclaimed experts speaking out at a conference against supposed chemtrails. Persuaded by this parade of seeming experts, she shared this video on Facebook. (For a sense of the theory's reach on this platform; at one point a chemtrails-themed Facebook Group had over 100,000 members.) Lois even wrote to her senator about chemtrails, but never received a response. By the time I met Lois in the fall of 2019, she was less focused on chemtrails, which had principally been her brother's concern — and she hoped an investigative journalist would someday expose the truth.

What Lois didn't know? In the intervening years, major tech platforms like Google had identified the chemtrails conspiracy theory and had begun to implement policies that had the eventual effect of leading to more fact-based and authoritative content rising to the top of the page of chemtrails searches. As of this writing in August 2020, for instance, if you conduct a YouTube search for "chemtrails," the first videos that come up are debunking videos rather than conspiracy videos. YouTube has also inserted a box at the top of the search linking to the Encyclopedia Britannica entry for "contrail"; this encyclopedia entry also debunks the chemtrails theory.

This recent change in Google policy allowed for an experiment. Our researcher asked Lois to go home and, over the next week, to re-open her investigation into chemtrails. At the end of the week, we called up Lois. The difference was remarkable. She said, "I found some new articles that debunked it. I'd have to say I'm not leaning towards not thinking chemtrails are real. I don't think they're spraying chemicals." Is there even such a thing as "chemtrails,"

as distinct from normal airplane contrails, we asked? "I'm leaning 80-90% no," Lois concluded.

What Lois's story demonstrates is that if technology platforms surface the right kinds of debunking content, it *can* have an effect on people who haven't yet become deeply attached to a conspiracy theory. In other words, through our ethnography we determined that it does seem possible for tech platforms to do what the debunker Mick West does, at scale.

What distinguished Lois from Jennifer is the relationship between what happens offline and online. A person who is heavily engaged in one and not the other is someone who can decrease or step back from a conspiratorial worldview. In Jennifer's case, her offline and online behaviors are melded together and influence each other. Her social engagements with other believers in the library revolved around online content that they dissect together as a group. Her relationship with her boyfriend is founded upon their shared belief in conspiracy theories. Her increasing isolation from her family and from mainstream sources such as Google are because of her belief in conspiracies – her parents can no longer relate to her and Google cannot be trusted. Lois, on the other hand, considered chemtrails but the rest of her life, offline, is not related to or motivated by a conspiratorial worldview. In fact, her research online was short-lived, brief, and kept private between herself and her brother. Lois is not part of a crusade or dedicated to finding the truth, in which chemtrails are linked to other conspiracy theories. This relationship between what happens on-and off-line was only discovered because of our ethnographic engagement with research participants. By visiting them wehre they live, we could observe economic changes to the town where they lived and appreciate why conspiracy theories could explain why some companies are so powerful and rich and their main street is shuttered. We could meet with their families, we could see what it means it to live off the grid, and we could witness the other parts of their lives that were not tied to conspiracies – their jobs, market investments, and church involvement. By observing both what happens on- and off-line, we could also observe what it meant to be a budding or a "light" conspiracist vs. a hardened, deeply entrenched and enmeshed conspiracist.

We also had demonstrated that a focus on treating some theories as harmful, and others as not, ultimately wasn't the most fruitful way to look at conspiracy theories. More fruitful was to be aware of the difference in types of conspiracy theorists themselves. Those who are newer to a conspiracy theory are the ones you are more likely to be able to reach with the facts, and therefore Alphabet's efforts to counter misinformation will be likeliest to have impact the further upstream they are. In other words, it's important to catch people at the top of the rabbit hole, *before* they really fall down it.

**CHANGING VIEWS ON CONSPIRACY THEORISTS**

We delivered our findings in a set of presentations for stakeholders across Jigsaw, Google, and YouTube in December of 2019. For many of these stakeholders, it was the first time they had encountered in-depth qualitative data about the lived experience of conspiracy theorists. Of course, many employees at Alphabet are highly specialized computer scientists and businesspeople; to bring an ethnographic perspective humanizing this segment of their user base was eye-opening. The study gave teams at Alphabet new language and perspective into how conspiracies work that they would not have had from other approaches.

Not long after the completion of this study, COVID-19 hit. Those who had been briefed on our research into conspiracy theories and its relationship to harm soon had to make difficult and rapid decisions about what sorts of COVID-19 content would and wouldn't be allowed on Google's platforms.

By April, YouTube had pulled thousands of conspiracy and misinformation videos related to coronavirus from the platform. It began surfacing an informational panel that linked to national health agencies' websites—like the CDC in the U.S. It also began aggressively enforcing medical misinformation policies around false COVID-19 cures, and it expanded that policy to bar promoting actions that go against recommendations from national health authorities. This expanded policy led YouTube to swiftly remove conspiratorial posts by Brazilian President Jair Bolsonaro, who had downplayed the virus.

Decision-making at an organization as large as Alphabet is diffuse, and it would be impossible to attribute these decisions to our ethnographic study alone. What we can say with confidence was that our study was a highly relevant and valued input to educate top decision makers of the harm of conspiracies at a crucial moment when Alphabet faced a flood of COVID-19 conspiracies.

We can be more precise and confident of our impact at Jigsaw itself, which after all was our direct client. Ethnographic research has been a core research stream at Jigsaw since its conception, but now its value is established and appreciated by the organization's top leadership, who participated in some of the conspiracy theory ethnography themselves. In June and July of 2020, ReD Associates and Jigsaw teamed up again, revisiting about half of our former conspiracy theorists, as well as a cohort of new ones, to learn what conspiracy beliefs they had about the COVID-19 pandemic. This time, we brought senior stakeholders not only from Jigsaw, but also from Google and YouTube's policy teams into the "field" (redefined as Zoom calls) with us. These senior Google stakeholders reported to us their hope that actually meeting conspiracy theorists would humanize and make more visceral their own understanding of the population their policies affect; "I hope to feel I understand them better than I would just by reading an article," one said. After their participation in fieldwork, these Alphabet policymakers confirmed to us that the interviews had achieved just that: humanizing an otherwise mysterious community.

## APPLICABILITY TO OTHER STUDIES

Based on our research experience, we have identified three learnings that others could consider implementing. First, given the sensitive nature of our topic—believers in white nationalism— and the negative connotation around being identified as a "conspiracy theorist," we adopted a multi-pronged approach to recruitment. We learned after engaging with recruitment agencies that it was too off-putting to recruit for white nationalists and that white nationalists themselves may not want to publicly identify themselves to recruiters. Instead, we identified a variety of proxies that could help us identify potential research participants. For example, one of the questions used in our recruitment screener solicited their news sources and we listed a mix of mainstream and conspiracy-specific publications. Once we identified people who regularly consumed conspiracy theory content, we held an initial conversation to gauge their level of familiarity with the types of conspiracies we were recruiting for and how frequent their engagement was in their day-to-day lives. Because this study was focused on a particular set of theories and we were searching for people who

might potentially act upon their beliefs, we selected those who were prolific researchers and could speak at length about their practices on-and offline related to their conspiratorial worldview. We also did not rely exclusively on recruiters. We engaged with people on social media who were self-identified conspiracy theorists and approached them for an interview once we connected with them. This required several conversations to explain who we were and what we were researching. We presented ourselves as social scientists interested in meeting people who believed in alternative narratives to the ones given in mainstream media. Given the fact that some believers in conspiracy theories are suspicious of others, we also relied upon our social networks to connect with friends of friends or acquaintances. Our social networks acted as a gatekeeper and facilitated these interactions that would not have taken place had we not had someone vouching for us.

Second, we benefited from having an expert join our team and help us analyze our data. We invited Mick West at the beginning of our research project and during our analysis phase. With his wealth of knowledge and experience, we were able to solicit his feedback on our initial insights and refine our models. Because we were working on a tight timeline, having a deep expert and debunking practitioner as part of the team provided us with frameworks in which to conceptualize the data and pushed us to think in more expansive ways around how to curb extreme conspiratorial worldviews and whom to target. We also recommend selecting an expert that the client is already familiar with and trusts because then the team can share the same references, concepts, and theories from the beginning and does not require the researchers to convince clients of the expert's relevance. In addition, it allowed the researchers to gain another perspective on how the clients view potential solutions to the challenge.

Third, throughout the project researchers worked closely with the client, especially the fieldwork. This is a common practice, but we draw attention to it because experiencing going to a remote location together and conducting the interviews together meant that the client already had an understanding of the everyday lives of conspiracy theorists. At the beginning of the project, we all had the notion that perhaps conspiracy theorists would be on the fringe, rather than mainstream – ranging from college-educated and working in large corporations to retirees. In addition, we came to the realization together that our initial hypothesis was wrong. We no longer believed that one theory was more harmful than another. Rather every type of conspiracy had the potential to be extreme, and that once an extreme version was adopted, it often indicated a way of seeing the world. We did not have to spend time convincing the client why our fieldwork had overturned this initial assumption and instead focus on telling a story that humanized, or rather, normalized conspiracy theorists. This portrayal transformed conspiracy theorists from being seen as "mentally ill" to a "regular" person. Helping technologists relate to their customer was the most helpful we could be as outside consultants.

We're optimistic that this more empathetic, holistic understanding of conspiracy theorists will be vital to decision making when they wrestle with refinements in their policies about how to handle disinformation, misinformation, and conspiracy theories on some of the world's largest tech platforms.

**Rebekah Park** is a Strategy Lead at Gemic. She has 20 years of experience as an applied anthropologist and holds a PhD in Anthropology from UCLA. She currently serves as a

board member of the Association of Legal and Political Anthropology of the American Anthropological Association.

**David Zax** is a Senior Consultant at ReD Associates, where he has focused on technology clients for the past three years. Prior to ReD, David was a technology journalist contributing to *Fast Company*, *Wired*, *The New York Times*, and other publications.

**Beth Goldberg** is a Research Program Manager at Jigsaw, where she oversees research on violent extremism. Jigsaw is a unit within Google that builds technology to tackle digital security challenges.


## NOTES

## REFERENCES CITED

Bartlett, Jamie, and Miller, Carl. 2010. The Power of Unreason: Conspiracy theories, extremism and counter-terrorism. London: Demos.

Brotherton, Robert, Christopher French, and Alan Pickering. 2013. "Measuring Belief in Conspiracy Theories: The Generic Conspiracist Beliefs Scale." Frontiers in Psychology 4: 279.

Coaston, Jane. 2018. "Why Conspiracy Theories Matter." Vox website. December 31. Accessed October 8, 2020. https://www.vox.com/2018/12/31/18144710/conspiracy-theories-trump-2018-qanon-soros-false-flags.

Dagnall, Neil, Kenneth Drinkwater, Andrew Parker, Andrew Donovan, and Megan Parton. 2015. " Conspiracy Theory and Cognitive Style: A Worldview." Frontiers in Psychology 6: 206.

Douglas, Karen M., Robbie M. Sutton, and Aleksandra Cichocka. 2017. "The Psychology of Conspiracy Theories." Current Directions in Psychological Science 26(6): 538-542.

Douglas, Karen M., Joseph E. Uscinski, Robbie M. Sutton, Aleksandra Chichoka, Turkay Nefes, Chee Siang Ang, and Farzin Deravi. 2019. "Understanding Conspiracy Theories." Political Psychology 40(1): 3-35.

Freeman, Daniel, and Richard Bentall. 2017. The Concomitants of Conspiracy Concerns. Social Psychiatry and Psychiatric Epidemiology 52(10): 595-604.

Graeupner, Damaris, and Alin Coman. 2017. The Dark Side of Meaning-making: How Social Exclusion Leads to Superstitious Thinking. Journal of Experimental Social Psychology 69: 218-222.

Grzesiak-Feldman, Monika. 2013. The Effect of High-Anxiety Situations on Conspiracy Thinking. Current Psychology 32: 100-118.

Fighting Conspiracy Theories Online – Park et al.

Hart, Joshua. 2018. "Profiling a Conspiracy Theorist: Why Some People Believe." Live Science website, September 26. Accessed October 9, 2020. https://www.livescience.com/63658-why-people-believe-conspiracy-theories.html.

Herrema, Martin. 2019. "How conspiracy theories affect low-level crime." University of Kent News Centre website, February 26. Accessed October 9, 2020. https://www.kent.ac.uk/news/society/21303/belief-in-conspiracy-theories-makes-people-more-likely-to-engage-in-low-level-crime.

Imhoff, Roland, and Lamberty, Karoline. 2017. "Too Special to be Duped: Need for Uniqueness Motivates Conspiracy Beliefs." European Journal of Social Psychology 47(6): 724-734.

Jolley, Daniel and Douglas, Karen. 2014. "The Social Consequences of Conspiracism: Exposure to Conspiracy Theories Decreases the Intention to Engage in Politics and to Reduce One's Carbon Footprint." British Journal of Psychology 105(1): 35-56.

Kluger, Jeffrey. 2017. "Why So Many People Believe Conspiracy Theories." Time website, October 15. Accessed October 9, 2020. https://time.com/4965093/conspiracy-theories-beliefs/.

McNamara, Audrey. 2019. "The Disturbing Link Between Conspiracy Theories and Petty Crime." Daily Beast website, February 26. Accessed October 9, 2020 https://www.thedailybeast.com/the-disturbing-link-between-conspiracy-theories-and-petty-crime.

Petersen, Michael Bang, Mathias Osmundsen, and Kevin Arceneaux. 2018. "The "Need for Chaos" and Motivations to Share Hostile Political Rumors." PsyArXiv. September 1. https://doi.org/10.31234/osf.io/6m4ts.

Preston, Elizabeth. 2019. "The Psychology and Allure of Conspiracy Theories." Undark website, February 27. Accessed October 9, 2020 https://undark.org/2019/02/27/the-psychology-and-allure-of-conspiracy-theories/.

Roose, Kevin. 2019. "YouTube Unleashed a Conspiracy Theory Boom. Can It Be Contained?" The New York Times website, February 19. Accessed October 9, 2020 https://www.nytimes.com/2019/02/19/technology/youtube-conspiracy-stars.html.

Stempel, Carl, Thomas Hargrove, and Guido Stempel III. 2007. "Media Use, Social Structure, and Belief in 9/11 Conspiracy Theories." Journalism and Mass Communication Quarterly 84(2): 353-372.

van Prooijen Jan-Willem and Karen M. Douglas. 2018. "Belief in Conspiracy Theories: Basic Principles of an Emerging Research Domain." European Journal of Social Psychology 48(7):897-908.

van Prooijen, Jan-Willem, and Nils Jostmann. 2013. "Belief in Conspiracy Theories: The Influence of Uncertainty and Perceived Morality." European Journal of Social Psychology 43(1): 109-115.

van Prooijen, Jan-Willem, and Mark van Vugt. 2018. "Conspiracy Theories: Evolved Functions and Psychological Mechanisms." Perspectives on Psychological Science 13(6): 770-788.

Whitson, Jennifer and Adam Galinsky. 2008. "Lacking Control Increases Illusory Pattern Perception." Science 322(5898):115-117.

Wood, Michael J. 2018. "Propagating and Debunking Conspiracy Theories on Twitter During the 2015-2016 Zika Virus." Cyberpsychology, Behavior, and Social Networking. 21(8): 485-490.

CATALYST

# Toxicity v. toxicity

## How Ethnography Can Inform Scalable Technical Solutions

JAMIE SHERMAN, PhD, *Intel Corporation*
ANNE PAGE MCCLARD, PhD, *McClard LLC*

*While a number of scholars have studied online communities, research on games has been mostly focused on the business, experience, and content of gameplay. Interactions between players within games has received less attention, and toxic behavior is a newer area of investigation in academia. Inquiry into toxicity in gaming is part of a larger body of literature and public interest emerging around disruptive and malicious social interactions online, cyberbullying, child-grooming, and extremist recruiting). Through our research we reaffirmed that toxicity in gaming is a problem at a global scale, but we also discovered that on a micro scale, what behavior gamers perceive as toxic, or how toxicity is enacted in gaming is different depending on cultural context amongst other things. The generalized problem at scale, and its particular manifestations on the micro level raise philosophical and technology design questions, which we address through examples from our own research and its applications in the industrial settings.*

*Keywords: Ethnographic research, Culture, Gender, Technology, Toxicity, Scale, Community, Global, Te*

## INTRODUCTION

Over the last year and a half, we—two anthropologists, one working for Intel, and the other a consultant—engaged on a variety of research topics related to digital game playing. We executed numerous multi-cultural ethnographic studies, each with a similar structure, and each building upon the previous studies. The basic design for each study included: 1) an online open-ended survey screening interview, 2) selection of participants based on stakeholder-defined criteria, 3) an introductory interview, 4) a week-long diary documenting daily life, gaming activities, frustrations encountered, social interactions, etc.; and 5) a follow-up interview. A majority of interviews were conducted in people's homes, but due to the COVID_19 pandemic, some were moved to online. In all, we conducted interviews with 49 participants across all studies.

The nominal focus for these studies was broad: to understand motivations, practices, needs, and "pain points" of different kinds of gamers in different places. We documented cultural themes around gender, social dynamics of teams, daily practices, consumption and engagement with game-related media, and attitudes and understandings of gaming as an activity and, for some, an identity. In parallel, and leveraging the "thickness" of ethnographic data, we asked questions and paid close attention to issues of "toxicity" or disruptive, unpleasant, and harassing behavior between players in gaming and game-related activities. This line of inquiry was motivated by a parallel project within the gaming team at Intel focused on how to reduce toxicity in gaming through AI moderated voice chat. That project—design of a real-time, scalable technical solution for managing toxicity in a game

voice chat system—is in progress as of this writing, fed in large part by the findings of our broader studies, with deeper investigation of voice chat interactions in games to come.

While a number of scholars, notably anthropologists, have looked at online communities (Boellstorff 2008; Nardi 2010), the study of games has, until relatively recently, been largely focused on the business, experience, and content of gameplay (Dyer-Witherford and de Peuter 2009). Interactions between players within games has received less attention, and toxic behavior is a relatively new area of investigation in academia, largely due to events like "Gamergate"(see Dewey 2014; Gray et. al. 2017). Inquiry into toxicity in gaming is part of a larger body of literature and public interest emerging around disruptive and malicious social interactions online, cyberbullying, child-grooming, and extremist recruiting (Adinolf and Turkay 2018, Fredman 2018).

Through our research we reaffirmed that toxicity in gaming is a problem at a global scale, but we also discovered that on a micro scale, what behavior gamers perceive as toxic, or how toxicity is enacted in gaming is different depending on cultural context amongst other things. We found that toxicity characterizations reflect the tensions, cultural beliefs, and attitudes within the local communities, and that these characterizations can even differ within the same geographic location between sub-groups. What gamers consider acceptable behavior in online voice chat varies dramatically by locale, by the game community of a particular game (e.g. PUBG, League of Legends, Overwatch, etc.), by type of player (e.g., competitive team player, recreational player), and even by ones gender, ethnicity, language, or race.

We raise three key questions in this paper, for which we do not promise definitive answers. However, we discuss each of them in the context of how our ethnographic work intersects and illuminates our technological pursuits:

1. What does scale mean? What is the relationship between the objectives of technological development at scale and the ethnographic project as an investigation into patterns of culture?
2. To what extent can ethnography's focus on the relationship between the local and the global facilitate generalizable recommendations for developing technical solutions? Is it possible to create technologies that are sensitive to local distinctions, and yet scale globally?
3. How does one go about identifying which aspects of a problem are important at a micro level, but not so important at a macro level? Is there a way to get to the core underlying problem and generalize it?

Throughout, we use examples from our own research, other's research on related topics, and refer to an ample body of anthropological theory that pertains to these questions.

**WHAT DOES SCALE MEAN?**

"Scale" has many meanings, so we begin our discussion with some of the different ways we think of it. One common use of scale from a linguistic perspective refers to the relative size of something as measured against some standard unit of measure—inches, feet, centimeters, number of people, and so forth. To scale a drawing, for example, means to keep the proportions the same when one is reducing a very large something to a smaller

something. All of the ways we present scale here pivot in some way on this general definition.

In the tech industry, one meaning ascribed to scalability in a technological solution is that it can serve the few or the many, technology that can grow with the needs of a business, for example. Another way of thinking about scalable technology is in how flexible it is, how easy it is to change, to add on to, or modify. And then, from a business perspective, scale is almost always about how many people will purchase or benefit from a technology—market size is of highest importance. Scalability of technology is important to the business because it increases market size ultimately.

Likewise, scale in research means different things. It can literally refer to the number of people included in a study, so a small-scale study might have as few as three people, and a large-scale study might include thousands to hundreds of thousands of people. Scale in research might also refer to the focus of study—micro or macro, local or global. All of these definitions of scale have relevance and implications in our work, and at times, the goals of scalability in each of these domains can be at odds with one another.

**Scale in Research**

Our job as ethnographic researchers in the tech industry is to generate new ideas, inform the design, and guide marketing of technologies that will appeal to the greatest number of people possible. Therefore, our research serves both notions of business scaling and technological scaling. Scale also comes into play for research itself, and how research is conducted, at what scale; some of the various methods that ethnographers use scale up better than others.

Participant observation is arguably the linchpin methodology of ethnographic inquiry, and is always present in ethnographic research; it requires deep dives into the daily lives of individuals, groups, and communities. Ethnographers, however, use a plethora of methodologies in addition to participant-observation, including, but not limited to, surveys, diaries, experimental frameworks, photo-elicitation, telemetry, and essentially any method that will help answer the questions that are in scope. Each method employed in ethnographic inquiry has different scaling characteristics.

By way of example, intensive participant observation research is not easily scaled to include as many participants as survey research. Not only are large-scale participant-observation studies onerous for the amount of data they generate, but they are prohibitively labor-intensive and too costly for most companies to justify executing them. This doesn't mean that they are not valuable; it just means that this key aspect of ethnographic research is usually done at a small scale. One question every practicing ethnographer gets is how does one determine whether ones findings are spurious given the typically small sample sizes of participant-observation research? There is not an easy answer to that question. The quality of the research depends on multiple factors, among them: how good the researchers are, how good the recruited participants are, how good the line of inquiry is, and how broad a net is cast. Fortunately, ethnographers practicing in industry do not rely solely on participant observation. We would argue that participant observation is the method with the most explanatory power of all of the methods we use, which is why it is so critical. Without it, and the deep insights it reveals about people's beliefs, views, and behaviors, other data in our tool box would be of limited value.

Let's consider survey research for a moment, one of the tools in an ethnographer's research kit. It is very scalable, and from a corporate perspective gives the best bang for the buck; it is relatively fast to collect the data, and per-participant costs are lower than for participant observation. One can ask as many or as few questions as one wants to, of as many or as few people are required for statistical purposes, and in theory, statistically significant findings imply generalizability and have predictive value. Large sample sizes, unfortunately, give business decision-makers a false sense of security in the validity of the data. Because, as with any research, the data is only as good as the researcher, how good the questions are, and how good the survey respondents are (which is a story for another day). The biggest problem with survey research is that taken on its own—without participant observation--it gives little to no insight into people's beliefs, views, and behaviors, because the responses are completely devoid of context and it is impossible to know anything about the thought processes that went into giving a particular response.

With all of this said, different business and technology related goals, require different research methods to achieve them, methods at different scales. For example, one cannot design a new product based on a large-scale market research segmentation. Products need to be designed with individual whole human beings in mind, with a detailed understanding of individual daily lives, motivations, situations, goals, beliefs, biases, etc., which is one area in which participant observation excels. The converse of this is that one cannot understand market sizing, price sensitivity, or build predictive models solely by doing participant observation. That said, participant observation and other qualitative methods used in ethnography are invaluable for properly framing large-scale quantitative studies.

## Scale from the Business Perspective

From a business perspective scaling almost always refers to growing the market. The question is always, "How do we get more customers to purchase our product"? At Intel, this is ultimately about selling more chips, whether direct to consumers who build their own PCs, or to "Original Equipment Makers" (OEMs) who make and sell laptops and desktops with Intel chips inside. Over the past decade or so, compute options have diversified with smartphones and tablets now capable of nearly everything one once needed a PC to accomplish. In addition, overall penetration of PCs in the overall market has increased, while the general compute needs of most users is met or exceeded by the devices they already own or have access to. Thus while the extent to which the PC is dead or dying, as some predict, is very much up for debate, the overall market for PCs is no longer seeing the kinds of growth it once did, and consumers now have a wider array of options.

High compute needs of scientific and enterprise data analytics aside, video game play remains one of the most visible spaces where users and game developers continue to push the boundaries of what current devices can do. In other words, it is an important growth segment in a market that is otherwise relatively lackluster and under threat. For that reason, Intel is especially interested in protecting and cultivating this market of enthusiast users who spend more on their systems and renew or upgrade their equipment more often than other consumers. In that context, it became important for the Intel gaming team to better understand the *scale* of the problem of toxicity in gaming as an impact on the market. They wanted to know the extent to which harassment and abuse lead to otherwise enthusiast gamers choosing other kinds of hobbies versus switching to, say, other kinds of games, while

still enjoying video games overall and continuing to invest in it as a hobby. That question was one we decided to tackle through a quantitative study of current and former gamers that is was informed by our ethnographic work, but not part of the discussion here.

In addition to enthusiast gamers, Intel has known for some time that nearly everyone who owns a PC of any kind plays games on them. One of the questions that Intel's gaming team was interested in, then, was how we might encourage these "casual" or "mainstream" gamers to become enthusiast gamers, and what kinds of barriers are there that we could do something about. Thus in early research we actually focused on these kinds of mainstream, non-enthusiast gamers, exploring how gaming fit into their everyday lives, the choices they were making in terms of whether and when to play, what to play, and what device or platform to play on. In addition, we asked a number of questions about their affective relationship to gaming, and their identification with gaming, as well as their experiences with, and concerns about toxicity and harassment. What we found in that work was that mainstream gamers, if they played with others, tended to limit their play to family and pre-existing friendships such that toxicity, which tends to occur between players who don't know each other from in-person contexts, was of less concern. However, we also found that mainstream gamers actively shied away from identifying as "gamers" in part because they didn't want to be seen as "one of those people." Thus while they reported fewer bad encounters and worried less about harassment, the overall negative association of gaming with toxicity and other kinds of bad behavior did in fact represent a barrier to gaming.

Getting back to "scale," from the business perspective, then, scale is best captured in terms of the number of purchasers, the frequency of purchase, and the relative value, or profit, in that purchase. From the business point of view, more is nearly always better: more people (size) buying more often (frequency) and buying more expensive models (in Intel parlance "upsell").

## SCALE IN ANTI-TOXICITY CHAT ALGORITHM

### Background: Toxicity in Gaming

Social anxiety around the content of video games cropped up as early as the 1970s, when the game Death Race was pulled off the shelves due to public outcry over its depiction of running over pedestrians for points, but it wasn't until 2014 or so that interactions between gamers came under significant discussion. Far from the first example of harassment of women in video games, Gamergate was nonetheless amongst the first to garner widespread attention. A disgruntled ex-boyfriend posted a host of accusations against an independent game developer accusing her of a range of things, including trading sex for positive reviews of her video game. The man's post sparked a coordinated campaign of harassment and intimidation that went viral, spreading from the woman herself to her known associates, and to other women writing and posting about video games. The harassment of these women included rape threats, death threats, coordinated email campaigns against websites, revealing photos emailed to employers and relatives, and public posting of home addresses and personal information in a tactic now commonly known as "doxing."

While Gamergate played out primarily across non-gaming social media platforms such as Twitter, YouTube, reddit, and 4chan, the notoriety of the campaign brought increased attention to the culture of sexism and racism in the video game industry, where women and

people of color are radically under-represented, systemically excluded, and routinely harassed. In addition, it brought heightened awareness to player interactions, where likewise women, people of color, LGBTQ+, and other minority communities are routinely and often quite viciously attacked (see here, here, and here, for just a few examples).

## A Scalable Technological Intervention

Against this backdrop, in early 2018, Intel drove a series of forecasting efforts focused on video game play. In that work, and in a range of subsequent studies including our own, toxicity between players emerged as a top concern for game players and was identified as a potential damper on market growth (limiting the scale of the market). The 2018 findings effectively shifted perceptions within the company from seeing player harassment in games as a moral and ethical problem facing individual players and game companies to perceiving player to player harassment as a business issue with potential implications for profit and sales, and a technical question regarding how and what we might "solve" using technological means.

While cognizant that the challenge is complicated, where harassment is frequently multi-modal and multi-platform, anti-toxicity work at Intel approaches the problem with a framework that is explicitly scaled, framed in terms of "crawl-walk-run." In this context, "crawl-walk-run" is an elastic concept applied both to the overall challenge of addressing social issues like this one by means of technical solutions, and more narrowly in terms of how a technical solution works by setting reachable goals ("minimum viable product") with the intention to add, grow, and improve over time.

From an overall domain perspective, team leadership focused in on voice chat as a key area for technical intervention. One of the (many) challenges in addressing toxicity is in identifying when and where it is happening. Today, most games rely on a combination of algorithms that screen text exchanges for transgressive vocabulary determined by individual game or organizational standards, and on users reporting of other user transgressions. The sheer volume of interactions, as well as the way that these interactions are fragmented across gaming and social media platforms make identifying incidents challenging. The Intel team began to focus on voice chat because, while text chatting is ubiquitous, there are existing tools for screening, and text itself leaves written records that game companies and moderators can review. By way of contrast, voice chat is transitory, the compute costs and privacy challenges of recording all exchanges prohibitive, and there are few to no existing tools. Here, the aspect of scale (or crawl in relation to walk and run) is about the focus on a relatively narrow context: screening voice chat interactions for "toxicity."

Within the voice chat anti-toxicity project, there are also internal notions of scale that are both additive and progressive. Initial efforts toward a "minimum viable project" focus primarily on the textual meanings of expressions and words identified by the algorithm in English (currently capturing UK and US based expressions). Future plans include areas for "scaling" such as the addition of community specific expressions, adding new game communities, adding new languages, and drawing on more vectors for analysis for better identification of the emotional tenor of an exchange – aspects such as tone, velocity, or decibel level have been mentioned.

Yet a third aspect of scale applies to the ways the team envisions the algorithm being used and by whom. Identifying offensive speech can be used by gamers to screen or silence

others, to document or report others for offensive behavior, and by gaming and game-related platforms to "triage" interactions for human moderation review, or to take automated action for specific kinds of transgression. The various ways the underlying capability can be applied, whether through unique applications or bundled into a single application with multiple features and stakeholders, represent an aspect of scale linked to, but not quite the same as, notions of scale embedded in the business perspective.

## LOCAL SENSIBILITIES V. SCALABLE TECHNOLOGY

This paper draws on a series of 4 in person and remote ethnographic studies conducted over the course of a year from March 2019 to March 2020. In addition to these studies, one of us conducted a series of 8 informal interviews with employees in the US who self-identified as enthusiast gamers, and fielded a quantitative survey with digital game players, non-players, and former players in the US to better understand overall patterns in leisure time practices and decisions to play or not play digital games, including the role of harassment or toxicity in games.

The first of the ethnographic studies in March of 2019 focused on "mainstream" and "casual" players in the Greater Los Angeles, California area. In that study we were particularly focused on better understanding the choices these players made across gaming devices, and the factors that made them more, and less likely to play overall. For the purposes of that study we defined "mainstream and casual" as players who spent less than 5 hours per week playing digital games of any kind on any device.

The next two studies, conducted in China and India looked at more enthusiast gamers, people who played at least 5 hours per week on average, and whose primary platform for gaming was either a PC or a console such as Microsoft Xbox or Sony Playstation, or both. In that study we were similarly interested in deeper understandings of motivations to play or not play, and in platform and device choices throughout the week.

The final study in this series was initially planned for Katowice, Poland, to coincide with a major e-sports event, and focused on competitive players and fans attending that event. In wake of the emergence of Covid-19 as a global health concern (not yet considered a pandemic at the end of February 2019), we pivoted to a mix of in person and online virtual meetings with participants in Portland Oregon, and several we had already recruited and who are based in the US, Europe, and Iceland. By the time we planned and executed this study, focus for the Intel gaming team had shifted toward Intel's "traditional" market of enthusiast and competitive gamers and this study was specifically designed to dive deeper into the specificity of competitive e-sports gamers.

Across these four studies, we spoke to a total of 49 video game players from diverse backgrounds, abilities, and levels of engagement with video game play in 5 countries. A few broad patterns emerged. Gamers of all kinds and abilities perceived toxicity and harassment as a critical problem in a guided question, part of the 7-day diary. In daily un-guided tracking of challenges and frustrations throughout the week, competitive gamers were far more likely to mention unpleasant encounters with other players. Mainstream and casual gamers who were less likely to play with people they did not know well in person did not cite toxicity or harassment in any of the sessions tracked in the study and were much less likely to mention personal experiences of that kind in the ethnographic and conversational interview portions of the studies.

In our interviews where toxicity emerged as a primary concern for participants, we were struck by all of the different ways gamers talked about and experienced toxicity, and the fact that what was considered toxic in one place was not considered toxic in another place. Furthermore, we found that even within a geographic location, that the very essence of toxic behavior could differ based on regionalism, language, or ethnic identity. Below, we tell three stories about toxicity in three countries—the US, China, and India. Each story illustrates a significant local sensibility about toxicity that we didn't see in the other countries. With the exposition of these stories, we challenge the reader, and the EPIC audience to think about the following questions:

1. To what extent can ethnography's focus on the relationship between the local and the global facilitate generalizable recommendations for developing technical solutions?
2. Is it possible to create technologies that are sensitive to local distinctions, and yet scale globally?

As a reminder, one of the primary directives from the company, was to assist in the development of the anti-toxicity chat algorithm we described before, a scalable technology that was being designed for scalable business purposes. We will discuss some of the thought processes and recommendations that have emerged based on our research.

## US STORY

Princess is an experienced competitive player in her late 20s who began playing video games with her father and at a young age. She has participated in competitive e-sports for several years as both a competitor and as coach of a local youth team. While the rampant sexism she encounters in gaming has made her more determined to play and to win, she describes it as exhausting and frustrating. Over time, she has developed tactics for avoiding toxic interactions directed at her. She uses a screen name that is non-gender specific. She generally avoids voice chat altogether, even in games where it is useful to the game, telling others that her microphone is broken (it's not). When she does use voice chat, she told us that when other players hear her, they frequently ask if she is "a girl or a squeaker" where squeaker means a pre-pubescent boy. Often, she said, she answers "squeaker" as she gets less hassled than if they think of her as a woman.

The way that Princess avoids voice chat was echoed by other women we spoke to in the United States, and men and women alike said that they prefer to use third party voice chat services such as TeamSpeak and Discord (today, primarily Discord) where they have more control over who is on the chat in lieu of in-game voice features. At the same time, using such platforms does not necessarily prevent players from experiencing toxicity, and in some games, in-game voice chat is an important element in the gameplay itself, forcing players to choose between exposure and success in the game.

In general, in the weighing of success vs the risk of exposure to offensive behavior, winning is frequently seen as more important, with toxicity as the price to pay. As Courtney, another female player in her 20s put it,

> it's one thing if you're talking offensively but you're still performing well; unfortunately, because of how toxic the overall community of League of Legends is, most people are like, eh, they said really racist stuff all game. But they contributed.

On the other hand, for Courtney, and others we spoke to, toxicity was experienced not just through language in speech or text, but through actions. She gave us the example of a team member who effectively sabotaged the match by actively helping the opposing team, or by failing to help the team. This kind of behavior, as she said "hits a lot harder" because losing a match has consequences not just for that day, but can also impact ones ranking and resources in the game overall.



Figure 1: In the US, sexism and racism are at the center of US players' experiences of toxic behavior in gaming. These experiences are tempered by perceptions of gaming overall, where some poor behavior is to be expected and tolerated.

While Courtney told us this story in the context of harassment she had experienced, for some users, match sabotage was a tactic used to enact revenge on harassers. Paula, a participant in a user feedback pilot focused on anti-toxicity told us that was her favorite response to encountering both racist and sexist harassment in gaming. She couldn't "just quit" she said, because quitting before the end of the match triggered consequences from the game, which could include reduced standing in match making and temporary bans, but that she would often just "stand back and let them die."

Courtney, it should be noted, is White. Gamers of color like Paula and also including several Intel colleagues who are African American and gamers themselves told me that they, like Princess, tended to stay off of voice chat, and to choose carefully who they played with. As with the women of color that Kishonna Gray studies, some of the them self-segregate (Gray 2012). For these women and African American players, qualities of voice were clear "tells" where speaking online over voice chat revealed them as female or non-white or both. Thus participation in voice chat was particularly fraught. Yet for those players who "passed" as Princess did, in saying she was a "squeaker," or Sam, a competitive e-sports player who is Asian, yet whose voice does not necessarily reveal his ethnic background, the default assumption of straight white male-ness, and the rampant use of ethnic and sexist language

could become deeply uncomfortable and cumulative over time, leaving players deeply ambivalent about those experiences. Sam, for example both downplayed the issue, saying "it's just part of the game" and that "people get really intense" but later said that people's anti-Black and anti-Asian comments would stay with him and really impacted his self-esteem and sense of self.

From these stories, it becomes clear that sexism and racism are at the center of US players' experiences of toxic behavior in gaming, but that at the same time, these experiences are tempered by perceptions of gaming overall, where some poor behavior is to be expected, and through tactics of avoidance (not using voice chat, not playing with strangers), revenge (letting them die), or, in the worst scenarios, quitting the game and finding a new game to play.

## INDIA STORY

Toxicity in gaming in India plays out quite differently than it does in the US or in China. One major difference in India relates to the gaming infrastructure, which is the most diverse of the three countries. While many of the communications platforms used in India are the same as for the US, players game on servers that are located in different regions of their own country, where there are significant ethnic and linguistic differences, and they also play on servers in neighboring countries like Pakistan and Singapore, countries with whom India has complex historical relations. This diversity in infrastructure gives rise to opportunities for toxicity that are not as prevalent in the US and China. The biases that were articulated related to cultural sensibilities of hierarchy in the Indian social system, and regional identities and linguistic markers. India also had the most hostile environment for female players.

One of our participants was a young woman in her 20s who got into gaming after meeting her boyfriend, who organizes local gaming meetups and competitions. For Divya, while she has had a number of bad experiences while gaming, those experiences are of a continuum with the overall sense of powerlessness and frustration linked to her position as a young woman more broadly, both online and off. In one story she told us, a player she didn't know cursed at her while playing PUB-G. Later, he joined a community linked to a different game and posted negative comments about her and her game play. He then left the community but seemed to return and then leave again, perhaps, Divya thought, to check on the effect of his postings. When she looked the other player up, she discovered he was a high ranking player in the game. She philosophized that perhaps, then, it wasn't so bad: a high ranking player had noticed her, and taken the time to seek out her community to comment on her play, and then returned later to see if she had replied. She brushes it off: "it comes to mind sometimes, but it doesn't affect me that much." In another story she told us, another player talked a lot about how bad she was. In that instance as well, she said, it was better to stay quiet because she didn't know that game as well as the other player, and also because she doesn't like to say much when she is angry. She told a similar story about her college life where, she said, she doesn't care much for her faculty, but when someone says something mean to her, she tries to distance herself. "I would like to say some things to them, but I can't because they are faculty." Instead she writes her anger down – about the faculty, about things that happen in the gaming community, and about how she can't go out at night for fear of her safety. For Divya, all of these frustrations are of a piece, shaped by her experience of being powerless in relation to other people or the overall context.

Figure 2. Competitive gaming tournament in India. As elsewhere in the world, competitive gaming is male dominated.

For Sushanth, an aspiring e-sports player in his 20s, male, the tensions are national and regional but not universal. As he said,

> *that actually depends on your luck. Like if you're really unlucky, you might encounter them every game you play... It's just random, there can be one game where one Singaporean hates Indians, whereas in the next game, they can be like, one more Singaporean who doesn't care at all and just wants to like play the game.*

In addition, these international encounters can have a positive resonance. For example when he told us about playing with Pakistani players. "I mean, you usually see Pakistan or India like big enemies, right, but when you actually play a game, those games actually turn out to be really good players who respect others and stuff."

For Sushanth, then, playing on regional servers with players from nearby countries is a matter of luck. In some cases he encounters negative perceptions of Indian players with repercussions from verbal abuse to throwing a match, whereas in others, international servers humanize players from "enemy" nations, people who just want to play the game and respect others.

As in the US, skill in the game makes a difference. Sushanth's positive perception of some Pakistani players is shaped by his experience of them as "really good."

## CHINA STORY

The infrastructure story in China is colored by the fact that many game players make a concerted effort to get to servers outside of the country, accomplished by using VPN. Some players prefer playing on foreign servers. One young man, Ren, has met many foreigners this

way, and even went so far as to fly to Germany to meet one of his online friends. The theme of traveling to other places in games was prevalent among other players, as well, but for most it was virtual.

Another key difference in China, and one that might present challenges in the design of a scalable chat-based algorithm for identifying toxic behavior is the fact that they use Chinese proprietary applications, like YY, QQ, and WeChat for gaming communications instead of Discord and Twitch apps, which are more widely used in the rest of the world. They also play a number of games that are modified from the original games that are available elsewhere. For example, they play Game for Peace, a version of PUBG that was developed specifically to appease Chinese censors, and is played on Android mobile devices.

In China "mean people" or "aggressive people" were cited as a serious problem, but people were reticent to talk in detail about their personal experiences with negative in-game interactions. Because participants were somewhat evasive, we feel that we don't have an adequate understanding beyond knowing that online toxicity *is* a problem. We know that cursing and "rage" behavior are prevalent, and we also know that gender-based interactions can be strained.

Most men talk about toxic behavior in a couple of ways—the use of foul language, which did not seem to bother them much, and bad actors, people who did not pull their weight, or who disrupted play in some other way. One young streamer, Jun, told a story about being blocked when the system detected that he was inadvertently teamed up with a known hacker.

> *The system detected that I was playing with a hacker. I did not know that I was playing with a hacker. I never cheated nor played with hackers. My friends are all working on that day and I was randomly matched to a group. Eventually, the webcast was reported and my account got blocked.*

He viewed the hacker as a bad actor, but wasn't bothered by other behaviors he encountered. Mostly, toxic behavior seemed like a non-issue for the Chinese men in our study, even though players reported it as a nuisance in their diaries.

Many of the women gamers that we spoke with initially told us that they did not experience online discrimination or sexual harassment, but some of the stories we heard suggested otherwise. One young woman, Juniper said that because men viewed women as weak players, males always sought to help and to protect females. This was especially true when she was younger and just started out. Later, when she became a strong player, this gave her extra advantages in the competition. She and others reported that sometimes men even pretended to be women in games just to get help from other male players. In describing an early experience in the gaming world, she told us how her brother served as her protector and guide.

Figure 3. Interviewing Juniper with her boyfriend in the background. How she talked about herself and her gaming abilities changed with him in the room.

*People[males] there [in the game] all regarded me as a little girl, so they took care of me. No matter if it was a certain task or something else, they would do favors for me. I felt happy.*

Juniper went on to become a successful and relatively high-ranking player in male dominated games, like World of Warcraft and League of Legends. While we were interviewing her, her boyfriend with whom she lived, came into the room. He had overheard her talking to us about her prowess and ranking. He scoffed and said, "You are not that good." He went on to tell us that men are innately better at playing computer games; he claimed they are genetically wired with faster reflexes. Juniper immediately acquiesced and was quick to say that she was not as good as he was. This in-person dynamic played out repeatedly in other stories we heard from women; they consistently said they were not very good.

Overall, most Chinese women players agreed that men were quite polite and solicitous of them in games. Most said that the condescension did not bother them; they accepted that this is the way of the world. Related to this is the role that game playing has in match making. Both men and women seek opportunities to meet and interact with potential mates in the games they play, and several told stories of how they met their partners through playing games. Sometimes these relationships went wrong, and other times unwanted advances by male players were problematic, but nobody wanted to share the details of these more contentious online experiences.

Most women players employ strategies for avoiding conflict of any sort with players of any gender. Rose, a young woman who enjoys a range of games from League of Legends to adventure games, said aggressive behavior by both women and men is common, usually in the form of cursing. She simply chooses to ignore aggressive behavior because she just wants to have fun.

*Girls are also quite aggressive now. No matter in life or in game, women also become very aggressive….We couldn't let those bad people affect us. I wouldn't even think about having quarrel with them. I met people having quarrel in the game. They bought the in-game trumpet to curse on all channels. I don't even know what the meaning is behind that. They're so angry and still spend that much money. 100 [yuen] for 10 trumpets, just to say 10 words. I don't know how much they would spend just to have a quarrel with someone.*

Mei, a young mother who enjoys playing games, but doesn't want to engage with hostile online game environments, limits her play to "people she knows," some from real life and others she has come to trust through playing on teams with them.

*Because in online games like this, too many strangers are too complicated, that is to say. Sometimes they play games to look for having sex.*

When pressed, Mei admitted to having had some negative experiences when she was younger, but she did not want to go into details. Mei tries, for the most part, to only play games with friends and friends of friends to avoid toxic online situations in games.

*For example, my friend plays a game and invites me to join a chat group. My friend has some friends in real life and some net friends in the chat group. When we become familiar, we can find that some people are in the same city, so we can make an offline meeting appointment. For example, we can have coffee in a bar or play games while gathering.*

It became clear that the general sentiment in China is that women should not be interested in playing competitive online games. One of the phrases that we became familiar with while doing our research in China was, "I am not a normal girl." Almost every woman we spoke to said something like this about herself. The implication was that "normal girls" should be interested in fashion, shopping, and K-pop, not in playing video games. Furthermore, most of the women we spoke with, even if they were quite good players in actuality, downplayed their abilities, especially if men were present during our interviews.

## COMMON GROUND

In thinking across these three cultures, we were able to find some common ground when it comes to toxicity, most of which have implications when it comes to thinking about scalable technical solutions.

Sexism is an issue in all three countries. That said, how it is expressed and how it is experienced are different from place to place. For example, in the US and India sexism is frequently expressed in verbal insults which could possibly be detected by our chat algorithm, but in China it is expressed in less easily identifiable actions— offers of help and in condescension. The generalized problem at sexism at a macro scale is difficult to address through a single technical solution because of the nuances that are revealed at the micro scale.

In all three countries, the relationship between skill differences and toxicity is present. Potentially toxic interactions arise when there is a skill mismatch among team members or competitors. If someone on your team is not skilled enough, or is refusing to fulfill their role in the game, it can destroy one's ranking. One's ranking in a game is not trivial, often the

product of significant investments of time and sometimes money as well. The desire to win at any cost gives rise to heightened emotions that lead to toxic behavior, often in the form of verbal abuse in voice chat, potentially identifiable with a speech recognition algorithm. However, there are caveats.

For example, there are exceptions to the skill rule that are also generalizable. Toxic behavior is more likely to emerge and be experienced as toxic with encounters between people of different skill levels who have relatively loose social ties. Thus, if one tends to play with people who are familiar, the emergence of highly toxic moments is less likely to occur, and also the toxic moments that do emerge are less likely to be perceived as toxic. It is the difference between your best friend calling you a [imagine an insult] when you make a winning move against her, and having perfect stranger calling you the same thing. Having a stranger hurl an insult at you is usually perceived as a worse transgression than having someone known to you do it. This is getting into some shaky areas for our algorithm. Can it learn who is known to you?

A correlate to the relationship of disparity in skill levels giving rise to toxic behaviors is that toxic behaviors are tolerated if the perpetrator's skill or relative social power in the team is high; people will tolerate bad behavior of the person who is doing the most to ensure a win. Again, we are unsure what this might mean for how a speech recognition algorithm might judge the toxicity of an incidence. Should it ignore potentially toxic speech acts of the highest ranked player?

Across all three geographies where we did research we observed that nearly all toxic interactions were described as a combination of speech, text, and other actions that crossed over multiple social platforms outside of the game itself. While an in-game or in-voice-chat algorithm might catch certain types of toxic behavior, this general finding reveals that the problem will not be entirely solved by such a technical solution.

Finally, universally, bad actors in games—spoilers and hackers—represent a particular type of toxic gaming behavior, another behavior for which there are no perfect technical solutions, and one that was beyond the purview of our work. Some games already have systems for automatically ferreting out such bad actors, but as we saw in the case of Ren in China, they can be quite heavy-handed, penalizing players who accidentally drafted a hacker into their team.

## THE PITFALLS OF OVER/UNDER GENERALIZING

Anthropology, as a discipline, has wrestled with the issue of generalization vs specificity since its beginnings as evidenced from theoretical and ethical positions that have swung widely from sweeping cross-cultural generalizations evident in early notions of evolution of cultures (Frazer 1890; Morgan 1877) and in direct comparisons between cultures (Mead 1928; Benedict 1934). Franz Boas, the "father" of American Anthropology favored the specific, although he argued for a historical comparative method (1940). Cultural relativism and the idea of ethnography as "thick description" have largely prevailed in the second half of the 20th century. However, in recent years, some anthropologists have argued that the lack of comparison has weakened the discipline (Borofsky 2019; Nader 2015).

Within the EPIC community, we have seen arguments arise regarding our ability to make meaningful cross-cultural comparisons, and about under or over generalizing. One thing is clear—it is paramount for businesses and institutions operating in a global market

place to design products and services that address both the general and the specific whether we are creating software and hardware solutions for automated driving (Rothmüller et al. 2018; Vinkuyzen and Cefkin 2016), trying to design technologies for domestic spaces (Pulman-Jones 2005; McClard and Dugan 2017), or trying to help our companies understand their customers (Anderson et al. 2017).

Our ethnographic research with gamers shows clear generalizable patterns across the three countries where we did research in the impact of abusive and harassing behaviors, including pervasive domains (e.g., gender, race, ethnicity), common behaviors (e.g., textual and verbal abuse, disruptive or hostile play tactics), and patterns of escalation where the most harmful behaviors transcend particular gaming episodes and platforms. The most troublesome behaviors are those that escalate to events on other social media platforms, expanding from individual hostilities to organized harassment, and from digital exchanges to actions in or affecting the physical safety, including doxing (publicizing personally identifiable information such as names and addresses) and swatting (sending emergency and other services to a physical address) in order to intimidate and harass.

By identifying these broad patterns, ethnography is instrumental in teasing out directions and trajectories for the movement from crawl to walk to run in the technical solution space. In other words, these patterns suggest not only relevant categories to identify, but suggest next steps for expanding or connecting to other kinds of solutions that, for example, trace behaviors across verbal exchanges and in-game actions, or from one game platform to another and to social media: solutions not available today, but important in defining how any given solution might fit into a larger picture.

At the same time, ethnography reveals levels of specificity that can feel overwhelming to technical and business teams, and clearly show the need for extensive localization that includes game specific and region specific language practices such but also goes beyond language to encompass region specific social tensions and common expressions and symbols related to those divisions. In addition, our interviews revealed distinctions across game genres and gameplay structures that impact the ways that harassment unfolds. For example, competitive team and match based games where players are often paired with strangers, stakes and emotions run high, and tempers flare. As related above, negative experiences tend to be most acute where social ties are weak. Role playing games, by way of contrast, tend to involve more sustained engagement over time, where players create and join guilds, playing together for weeks, months, even years. In those games, toxicity and harassment stories focus less on sudden and isolated outbursts, but can involve longer term targeting and collective actions to isolate, bully, and intimidate victims.

Finally, the relationship between local and global and the frequent mixing of known and unknown participants, strong and weak social ties raise important issues of power, ethics, and responsibility. Specifically, who decides what is, and is not "toxic?" Today, this is left to individual gaming platforms to decide, and frequently up to players themselves to report. However, when considering the idea of technology that might screen interactions for offensive behavior, many players worry about the distinction between "trash-talk" which can be fun between friends, and harassment or trolling which is distinctly not fun for the victim, though perpetrators may perceive and represent it as "play." At the same time, while participants themselves may be comfortable with certain kinds of biased (sexist, nationalist, racist) expression amongst themselves, does that mean it should be allowable? If sexism

expressed in games takes place in contexts where men and women alike share a perception of male and female roles and abilities as unequal – is it toxic?

## OUTCOMES AND CONCLUSIONS

Sometimes the specificity of ethnographic research is at odds with telling a simple generalizable story, and our insistence on telling complex ethnographic stories may make it seem immediately more difficult to act from a developer and business perspective. However, the complexity of ethnographic research findings is what ultimately gives direction for evolution of a product by emphasizing the local experiences and social and cultural differences that exist in disparate markets. In other words, ethnographic research has the power help to create a roadmap for scaling a solution. So while an MVP may barely scratch the surface in the "Crawl" stage, later implementation will have the guidance embodied in the general and specific findings gleaned from research.

How has the complexity gleaned from our ethnographic research on gaming informed scaling of the anti-toxicity voice chat algorithm? We showed the development team that localization, not just language translation, matter. We showed them that regional tensions, social hierarchies, vocabulary, and gaming community cultures matter too. For the "minimum viable product" (MVP) the team is working on an algorithm that accurately interprets a range of vocal intonations and accents in American English, and they now have roadmap that includes local sensitivity.

To wrap up, while, participant-observation is the cornerstone of ethnographic research, it is not the only hammer in our toolkit, and using multiple methods across a research domain enables researchers to make impactful contributions that reveal and bridge between micro and macro perspectives. Participant-observation helps drive the development of scalable technical solutions by revealing underlying patterns and common ground, by teasing out the specificity of local interpretations and enactments. The ethnographic research we discussed here has formed a critical part of developing an effective strategic roadmap for meeting users where they live and play.

**Anne Page McClard** holds a doctorate in cultural anthropology, and has worked in the technology industry for more than 30 years. Anne uses ethnographic research to influence and drive product design and strategy, in both consumer and B2B markets. Throughout her career, she has sustained an interest in gender issues in academia and technical industries.

**Jamie Sherman** is a cultural anthropologist (Ph.D. Princeton. 2011) and research scientist at Intel Corporation. Her research background is in techniques and technologies of self-transformation, performance, and dynamics of race, gender, and play. Since joining Intel in 2012, her work has focused on emergent technological practices from quantified self, to live motion capture projection. Her current research develops usages and drives strategies for video game play, media creation, and online toxicity.

# REFERENCES CITED

Adinolf, S. and S. Turkay
2018    "Toxic Behaviors in Esports Games: Player Perceptions and Coping Strategies." In *Proceedings of the 2018 Annual Symposium on Computer-Human Interaction in Play Companion Extended Abstracts (CHI PLAY '18 Extended Abstracts)*, pp. 365–372. DOI:https://doi.org/10.1145/3270316.3271545

Anderson et al.
2017    "Creating a Creators' Market: How Ethnography Gave Intel a New Perspective On Digital Content Creators," *Ethnographic Praxis in Industry Conference Proceedings,* pp. 425-443 .

Benedict, R.
1934    *Patterns of culture*. Boston: Houghton Mifflin Company.

Boas, Franz.
1940.    The Limitations of the Comparative Method of Anthropology. In Race Language, and Culture. Pp. 270-280. New York: Macmillan.

Boellstorff, T.
2008    *Coming of Age in Second Life: An Anthropologist Explores the Virtually Human*. Princeton: Princeton University Press.

Borofsky, R.
2019    Where Have the Comparisons Gone? *Society for Cultural Anthropology*.

Dewey, C.
2014    "The only guide to Gamergate you will ever need to read" *Washington Post*. https://www.washingtonpost.com/news/the-intersect/wp/2014/10/14/the-only-guide-to-gamergate-you-will-ever-need-to-read/

Dyer-Witherford, N. and G. de Peuter
2009    *Games of Empire: Global Capitalism and Video Games*. Minneapolis: University of Minnesota Press

Frazer, J. G.
1890    *The Golden bough, a study in comparative religion*. London: Macmillan.

Fredman, L.A.
2018    *Not Just a Game: Sexual Toxicity in Online Gaming Hurts Women*. Unpublished Dissertation, University of Texas at Austin. http://dx.doi.org/10.26153/tsw/2112

Gray, K. L.
2012    "Intersecting Oppressions and Online Communities," in *Information, Communication & Society*, 15:3, 411-428, DOI: 10.1080/1369118X.2011.642401

Gray, K. L., B. Buyukozturk, and Z. G. Hill
2017    "Blurring the boundaries: Using Gamergate to examine 'real' and symbolic violence against women in contemporary gaming culture" in *Sociology Compass*. 11:12458. https://doi.org/10.1111/soc4.12458

McClard, A. and T. Dugan
2017    "It's Not Childs' Play: Changing Corporate Narratives Through Ethnography," *Ethnographic Praxis in Industry Conference Proceedings.*

Mead, M.
1928    *Coming of age in Samoa.* New York: William Morrow & Company

Nader, L.
2015    *What the Rest Think of the West : Since 600 AD.* Oakland: University of California Press.

Nardi, B. A.
2010    *My Life as a Night Elf Priest: An Anthropological Account of World of Warcraft.* Ann Arbor: University of Michigan Press.

Pulman-Jones, S.
2005    "Using Photographic Data to Build a Large-scale Global Comparative Visual Ethnography of Domestic Spaces: Can a Limited Data Set Capture the Complexities of 'Sociality'?" EPIC *Ethnographic Praxis in Industry Conference Proceeding*, pp. 128-139.

Rothmüller, M. et al.
2018    "Designing for Interactions with Automated Vehicles: Ethnography at the Boundary of Quantitative-Data-Driven Disciplines" *Ethnographic Praxis in Industry Conference Proceedings*, pp. 482-517.

Vinkuyzen, E. and M. Cefkin
2016    "Developing Socially Acceptable Autonomous Vehicles," *Ethnographic Praxis in Industry Conference Proceedings*, p. 522–534.

# Impact at Human Scale

Enacting change in the world is fraught with obstacles. How can individuals have a meaningful impact at a global scale? How do institutions, with their own decision-making logics and biases, impact individuals?

Session Curators: Nick Agafonoff, Lindsay Ferris, Tabitha Steager, Erin Taylor

# Scaling Through Meaning to Action

## What the Australian Bushfires Taught Me about Ethnography

CHARLIE COCHRANE, *Jump the Fence*

*This PechaKucha gives a personal perspective on the ethical dilemmas around the impact of an individual's actions, and the meaning of an ethnographer's projects in the context of the scale where these play out. The story begins with the spectacle of the 2020 Australian bushfires and reflects on their enormous scale. Within this context what is the meaning of individual actions to limit global warming?*

*The story shifts to the work context and explores the dichotomy of human impacts vs. the marketing metrics that typically measure success. Using an example of a research project with an overtly purposeful aim we explore the tension between ethnography as a tool for understanding the problem and the question of whether the scaled result truly addresses the end-users' problem.*

*Returning to the bushfires, we again look at the scaled government response and the question of how successfully this met the needs of those impacted. We explore the different ways that initiatives can be scaled and recognize that smaller initiatives, tuned to end-users' needs may be those that produce the most human impact. We conclude with the challenge to expand the remit of ethnography beyond problem diagnosis through to end-results.*

Aftermath of the Currowan bushfire © by photographer Stephen Dupont

**Charlie Cochrane** (charlie@jumpthefence.com) directs an ethnographic and qualitative research consultancy firm in Australia. He has over 25 years experience in ethnographic and qualitative research, marketing and advertising in both the UK and Australia. He has an MA in anthropology from Cambridge University and has trained in psychology and psychotherapy for professionals at Spectrum in London. He pioneered ethnographic research in Australia. He has conducted projects for government, service organisations and blue chip manufacturing companies in both the UK and Australia for household name companies such as Nescafe, Campbells, American Express, and GIO as well as projects for state and federal government. He also teaches an ethnography unit to MBA students at NTU in Singapore.

# Sustainability

## Addressing Global Issues at a Human Scale

LEE RYAN, *More Things Considered*
LOUISA WOOD, *More Things Considered*

*Our tiny provocation is that the word "sustainability" is not sustainable. Just using it is sabotaging our efforts to build a better future for the planet. Despite decades of global sustainability discourse, the world is still going to hell. What's gone wrong? Our paper is about willful ignorance and complicity at a global scale; the benefits of small talk; and a better, more effective word than sustainability.*

## INTRODUCTION

> '"Sustainability" is the dream of passing a liveable earth to future generations, human and nonhuman. The term is also used to cover up destructive practices, and this use has become so prevalent that the word most often makes me laugh and cry.' Anna Tsing (2017, 51)

We (ethnographers working with organisations) recognise the importance of sustainability, defined by Hunter as "treating the world as though we plan to stay" (2020). Some of us work alongside organisations with far-reaching sustainability ambitions. The panel 'Agency & the Climate Emergency' at the 2019 EPIC conference asked:

> What is the ethnographers' role in dealing with a catastrophic climate crisis? Should we be exploring people's experiences of change, trying to use our insights to help drive individual and collective action at scale through organisations, or helping civil society deal with the consequences?

The problem this panel references is sustainability can only be addressed at a global scale, but the complexity of individual practice can only be understood at a human scale. To affect sustainability at scale, organisations need to tap into the practices and beliefs of individuals collectively shaping what's possible.

We have worked on multiple projects where our clients wanted to understand or communicate sustainability. In 2019 we worked on a global qualitative project at the intersection of sustainability, food and young people. It involved conducting research in six cities and spending time with people, online and in person. The research included market tours curated by local teams to experience how young people in a city buy and eat sustainably. We had conversations with experts in different cities, from academics who specialise in the field of sustainability, to chefs, agency folk who develop different types of communications, packaging specialists, people looking at emerging business models in social innovation and those who lead significant sustainability efforts within global organisations.

While we identified valuable insights that addressed specific questions for the client, we ourselves finished in 2019 with more questions than answers.

This paper discusses the questions this project provoked, and covers in detail our 2020 research response. In contrast to the global project with multiple phases and cities, we decided to go small. Our project in Auckland in early 2020 involved eight participants. We asked each participant three questions relating to sustainability, without actually mentioning the word 'sustainability.' This approach gained meaningful insights. We will cover three strands emerging from this study. First, we discuss the ongoing presence of silence in sustainability and the slightly provocative question of whether the word 'sustainability' is essentially meaningless. Should we continue to use the word 'sustainability' if it means so little and is largely empty? Second, we discuss small talk versus big talk and the mismatch between how organisations and individuals talk. What could happen if we bridge that gap? Finally, we turn to the emerging theme of survivability. What does it mean if people are talking survival whilst there is 'bigger talk' about sustainability?

## WE ALL TALK ABOUT SUSTAINABILITY AS IF IT'S A THING

### Field Vignette: 1

*I'm sitting in front of my laptop. I'm opening up day 2 on dscout. It's the first market. There's the relief in finishing a discussion guide, wrestled over between you, your local team and the client, and then the tension of it landing in the messy lives of humans. Are they the right questions? Are they worded well? Have we connected with people enough? Today people are sharing videos of us of what's in their fridge and pantry. I start watching. There's the familiar pleasure in hearing people narrate what they see in their kitchen. At the end of the first video, I notice an awkwardness or silence. We had asked people to finish by showing us the most sustainable item in their fridge or pantry. The person paused. For a long time compared to the last two minutes. "I guess this" as they landed on a bottle of water. I watched other videos. Other people also were silent, hesitant and even reluctant to name anything in their fridge or pantry 'sustainable'. What made it more intriguing was they had been recruited for behaviours that the client had identified as relevant, sustainable behaviours so as to spend time with people already practising sustainability in some way.*

In 2019 we were struck how differently people conceptualise and experience sustainability. People inside organisations and agencies talk as if there is a tangible and universally accepted understanding of sustainability that is actionable at scale through products, policies and practices. The embedded organisational talk ('sustainable practices', 'sustainability creative agency', 'making sustainability', 'achieving true sustainability', 'the influence of sustainability') assumes a shared understanding. Individuals talk at a human scale across a diverse range of small, personal practices (recycling, plant-based diet, shopping local, workers' rights) that enable them to feel like they're doing something "good".

There appeared to be language gaps. If we introduced sustainability into discussion, yes, people would politely talk about it, and respond to examples provided of possible candidates for sustainable packaging or communications. But people did not talk comfortably, or even consistently, about sustainability. My sustainability is not necessarily another individual's or organisation's sustainability. Interpretations could be anchored in economic concerns (viability in business or in wages); social concerns (my family or community); environmental concerns (anything from the degradation of my local environment to global warming); or any elements within the means of production.

Our interviewees struggled to define sustainability and to produce tangible examples in

the context of their lives. The *Better Futures Report* by Colmar Brunton New Zealand (2019) is an annual survey focusing on consumer behaviour towards socially, environmentally and economically responsible brands. The survey found seven out of ten respondents were unable to name a brand leader in sustainability. When asked to pick from a list, sustainability credentials seem to align most closely with overall brand communications and advertising spend. Does this lack of recognition, even after decades of promotion, indicate the word sustainability has limited commercial value and meaning?

The finding contrasted with the experts we spoke to who had roles inside organisations who talked about sustainability as a knowable thing. The experts might allude to the 'intention action gap' around sustainability but not that the word 'sustainability' has a language problem. We detected a 'consumer' silence and also a type of organisational silence. The global brief requested we compile global and market-specific commercial research reports on food, sustainability and millennials for our client. As we spent time with people in different cities, we returned to the commercial reports to see how they framed and managed this language gap. Did they talk about the gaps in language? Trend reports were breathless in their announcement of shifts and changes, including changes in consumer activity in supermarket aisles and social media, but most reported attitude shifts captured in response to statements in surveys. The subheading of the 2018 Nielsen Report *The Evolution of the Sustainability Mindset* is, tellingly, *Get With The Program: Consumers Demand Sustainability'*. This trends report claims sustainability has increasing value for consumers but is not translating into consumers discussing sustainability. Can actions and discourse, therefore, be captured by the word 'sustainability' anymore?

A global project gives you the time and space to ruminate on significant issues. When we spoke with experts inside organisations, such as the Head of Sustainability at a corporation or university, they were comfortable discussing sustainability. This stood in contrast with 'regular' people we interviewed who mostly struggled to talk about sustainability, but what they did talk about was smaller behaviours. There were noticeable silences when we gently questioned organisations' assumptions about the inherent good and usefulness of sustainability, the language gaps, and what we miss in the present if the brand discourse assumes a particular future (bright, good, optimistic). What are we not discussing? Is sustainability a useful concept-metaphor? Moore (2004) claims concept-metaphors are valuable because they can maintain ambiguity and a productive tension between universal claims and specific historical contexts, as well as acting as a mechanism for all of us to communicate. Is sustainability valuable in this sense? Or is it worthy, but so vague and encompassing so many different associations that it is an essentially meaningless term? Further in her article, Moore cites Appadurai's use of 'scape' as a "space for action and thought not only for anthropologists, but also for… families and individuals." (2004,79). We have observed that sustainability can be a space for thought for anthropologists (based on books that continue to be published), but little evidence that it is a space for thought for families or individuals. What if "there is no there *there*" and by introducing sustainability as a topic into conversation we are introducing an artificial construct? We are, in effect, superimposing a concept on the field, rather than understanding how families and individuals might make sense of the conceptual space.

We aren't, of course, the only people raising issues around the language of sustainability. Hasbrouck and Scull, in *Hook to Plate Social Entrepreneurship: An Ethnographic Approach to*

*Driving Sustainable Change in the Global Fishing Industry* comment on the use of the word in the seafood industry:

> The imprecision surrounding the definition of "sustainability" has been passed on to consumers, who are largely confused about what it is that "sustainability" means when it comes to fish. "Sustainability," as a food label term, stands out in its ambiguity even among other ethical food choice labels, whose names are varyingly self-explanatory, such as "free-range," "shade-grown," "cruelty free," and "fair trade." (2014, 471)

Mike Youngblood uses a textbook definition in his introduction to the *Sustainability & Ethnography in Business Series*, stating that "Sustainability is an approach to acting in the world in a way that consumes resources and produces waste at a rate that could be continued indefinitely" (2016, 1). If we pause on that definition, what if sustainability as a concept does not connect sufficiently to how people are acting in the world in relation to their consumption or their waste?

What could it mean to work at a human scale? In her chapter 'Design Ethnography, Public Policy, & Public Services: Rendering Collective Issues Doable & at Human Scale', Kimbell argues the value of ethnographic practices is:

> not that they are human-centred but rather they provide a way to understand sociomaterial assemblages involving complex political, financial, social, and technological systems at human scale...even given limited resources, the analytical orientation of ethnography is productive for asking different questions and provoking new thinking. (2014, 163)

What would it look like to closely attend to human experiences and how people interpret them—to be in the room in a tiny way? What if people did not know that sustainability was on the agenda? What if we let the participants define the challenge—letting the language and concepts emerge from what participants experience and talk about as important and "remain open to what is actually happening" (Glaser 1978, 3)?

At the start of 2020, we recruited eight people across Auckland and spent time with each of them in their homes. (The fieldwork took place between the end of January and the start of February. At the time, COVID-19 was present but unnoticed in New Zealand.) Rather than ask the questions we wanted answers to directly, we let people talk about what they wanted to talk about. Practically, we asked three questions. We asked them to talk a bit about themselves; we explored what was currently on their minds, and finished up with discussing the future. We used listening, probing, and replaying to explore what emerged.

## SMALL TALK & LITTLE THINGS VS BIG TALK & GRANDIOSE GESTURES

We learned more about sustainability, both conceptualised and actioned, by asking these eight people in one city three questions unrelated to sustainability than asking a hundred people living in six cities across three continents around fifty questions that were mostly about sustainability. Our exchanges with the former group were of a safe, inconsequential,

prosaic kind that happens at the dinner table: "what do you do", the weather, and how we all know the host. This is small talk. Our exchanges with the latter group concerned the significant, extraordinary and far-reaching type that happen at the boardroom table: what are the biggest issues facing us today, tell me about your sustainable actions, and how about global impact. This is big talk.

We started each conversation by asking them to tell us how they would introduce themselves to someone they just met at a party. At some point, the second question was a version of what's worrying you, and if something is keeping you awake at night, what would it be? Because we only used these questions, we could relax into the conversation and to give people the experience of being listened to. The conversations had a meandering structure with neither party knowing the direction, structure or even topic.

As people talked about what concerned them, talk often naturally drifted to concerns about the planet. To give this some small, local context: it was the end of the Southern Hemisphere summer. We had all just experienced the sky going an unusual dark orange due to the Australian bushfires, which had disoriented people sufficiently that the police had requested that Aucklanders not ring 111 to report on the condition of the sky. People who wanted to swim or fish at beaches needed to consult different sources of information due to pollution impacting where you could go. At the start of summer, the David Attenborough documentary 'Our Planet' had screened. It included a scene where, as a result of a lack of sea ice in the Russian Arctic, hundreds of walruses drag themselves up a cliff and then, unable to get back down, fall to their deaths on the rocks below. The NZ government had banned single-use plastic shopping bags a few months earlier.

Across the eight people, we heard versions of what we now term 'small talk.' People specifically mentioned 'little conversations' they had had with others. One of our interviewees, Colin, recounted how he and his kids had recently talked about conversations they had when his kids were living at home:

> So I talked to my kids and one of the things we always talk about is how we had really good talks in the car, right? They're like Yeah, we did right you know, we just, well that song sounds pretty cool. I like their way right now. It's what is this rubbish? And you know, you'd be talking shite, but then you may talk about something else. So I know this boy in my class, likes this song and it's like, Oh, okay. Right What boy? Yeah, yeah. Okay.

Colin is talking 'shite' with his kids but the kids are able through the small talk to talk about the big things in their lives, and Colin is learning about the big things in their lives through the small talk in the car. The fact his kids are now in their twenties, and they remember it, and he remembers it, is quite meaningful. It's a conversation. Both parties are talking, listening and reacting to the environment they are in. We noticed an opposition between small talk and big talk. Sustainability is an abstract word, mostly used by organisations conveying a big message. We wonder what would have to happen for the large scale monologue and small scale conversations to be joined up.

The business press has promoted 'purpose' as a key asset for organisations in the last decade, in particular, drawing upon Sinek's book *Start With Why: How Great Leaders Inspire Everyone to Take Action* (2011) and Stengel's *Grow: How Ideals Power Growth and Profit at the World's Greatest Companies* (2011). This use of sustainability within organisations, alongside a changing climate, provides a context for renewed focus on big ideals articulated as 'purpose.'

The ESG (environmental, social, and governance) boom in investment (where since 2012, total assets in sustainable investing have more than doubled) has been explained by the publisher Visual Capitalist, who states that:

> With a wide range of global sustainability challenges and complex risks on the rise, investors are starting to re-evaluate traditional portfolio approaches. Today, many investors want their money to align with a higher purpose beyond profit. (Visual Capitalist 2020)

Marketing and PR weave sustainability into their reporting, their packaging and statements about new and upcoming products. Oatly released its first sustainability report in 2017, including the statement "Sustainability, that's what our business is." as the headline for the lead article in the report, which is an interview with the CEO. This is detailed with:

> We are not driven primarily by going to work and earning money. We want to make the world a better place, primarily by contributing something to society, and then making money. But sustainability—that's what our business is." (2017, 6)

H&M promotes its Conscious Collection, "Toyota pursues the creation of a sustainable society through its CSR activities", and BrewDog committed to becoming "the world's most sustainable drinks company." And "Unilever has been a purpose-driven company from its origins. Today, our purpose is simple but clear—to make sustainable living commonplace."

This contrasts with individuals who are reluctant to be the sustainability 'poster child'. Kerrin, one of our participants, recycles, farms worms, and sews blankets as part of a community group. She is reluctant to be seen as an example of something good, partly from guilt, and her sense of not doing enough.

> I guess, in a sense, with the worm farm, obviously, reducing food waste, what's going into rubbish, things like that. I guess recycling clothes and sometimes I'll purchase from an op shop [charity shop]. But mostly it's crockery and just having a knack for old things, I think, yeah, like with my sewing group, we use donated stuff But I guess I personally would not be a spokesperson for it because I'm terrible.

Later in our conversation, Kerrin discusses companies in the fashion industry, in which she works. These companies are discussing their impact on their environment, but from her perspective are still not doing enough as they fly clothing in from other countries.

We could contrast Kerrin with the more angst-ridden Paige, who is torn about her continuation to eat meat. She does other things: walking her child to school, picking up litter, recycling clothes. She sees these little things as being at odds with grand gestures.

> No, yeah, I haven't done any grand gestures. Like I'm aware of all the meat and how much meat industry uses. But I still love meat. So that won't change. my cousin and friends vegan. Yeah. So we have these discussions a lot. … Just when I see her because you know, what's worse being aware of, you know, the meat industry and all that stuff, being aware of the effects but still eating meat. Or just you know what I mean? Like, what is worse? Is it worse, but I'm aware, but I still eat meat and I still do this. But it's just delicious. That's just my choice at this moment.

Organisations use the term 'sustainability' in claims they make, though as noted recently in the FT, "consultants say that most still tend to focus on applying their existing initiatives to conform with the UN's [sustainability] goals, rather than on reshaping their business models to make them sustainable" (2020). Our participants shared similar scepticism around organisations and sustainability. When discussing the communications team within her organisation tasked with promoting sustainability, Talia said:

> I think that's only because of trends… putting out posters about what you can do to be more sustainable… She's great at what she does but if she was practiced what she preached she wouldn't she probably wouldn't do x y & z …[They are] tapping into the environmental hashtag…we care….

The look on her face, and tone in her voice when mentioning the hashtag, were amplified through finger movements indicating cynicism. Talia described how her organisation publishes posters on sustainability that are put up in key work and public areas as a tick box exercise. We discussed how, once it becomes a communications exercise, it is no longer a real thing.

Our participants noticed the difference between how organisations talked about sustainability and the big things the organisations do to combat climate change. This contrasts with non-organisational small talk. The small things are doable, affordable, and practicable. These make people feel validated about trying their best to make their [local] small world a little better. People are doing small things, but it's about doing fewer bad things than doing good things per se. Underlying this was guilt and a suspicion that none of this is actually good or making a difference.

## BIG TALK IN SMALL CIRCLES, SMALL ACTIONS AS THE LITTLE THINGS ARE IMPORTANT

People talk about doing "*little things*" by using small talk to talk about bigger issues. This approach is not didactic, not direct, but rather uses oblique, everyday chat and small actions. Kerrin, our young twenty-something who had a worm farm stated:

> I guess I might seem a bit odd to some people when this is real big issues out there. I guess I feel like I'm just gonna go and do all these little things. I guess in my view, little things mattered to me, besides just how I was raised, so if I can contribute even in a small way, that makes me feel better.

People exchange within small inner circle(s), talking with and listening to each other. These are genuine conversations. Organisations communicate through monologues. Brands in this space typically talk in big ways, about big things. They broadcast through advertising, packaging and annual reports. People work in their own way to acknowledge the climate is changing. In practice, this looks similar to observing the local effect of climate change; they are no longer able to fish at their local beach. It's not about the whole world, but their world. It's talking to their kids about recycling. Emma talked with us about the climate as a wider issue that concerned her, and how recycling was something she was conscientious of. She discussed how her son, 22 years old and sharing a house, he had rung her agitated at his

flatmates' poor recycling practices. She commented:

> That's quite interesting that now I can see all the years of just little conversations I
> had with them like it has actually filtered through and then it's quite cool. Things I
> would have never expected that he would worry about like recycling especially
> being a young man as well.

People feel safe with immediate family; they lean in with those conversations and integrate small talk and small actions. They try to be encouraging. They take small actions to make themselves feel better or actually just less bad and less guilty. Emma went on to tell us, "you know, if I, if we're out you know, I've got recycling I'll bring it home I won't just throw it in a bin while we're out I'll bring it home and put it in my recycling. Just a little thing that I feel like I'm doing that's better." How we treat each other informs the social and natural environment we live in.

## OUR BIG TALK IS IN OUR SMALL CIRCLES; OUTSIDE IS RISKY—WE ARE QUIET

### Field Vignette: 2

*I'm looking down at my field notes as we debrief in the café. Words like 'minefield' and 'crazy' had been mentioned by our roofer who had bought an electric car for social encounters. We had recently spent time with a woman who with her husband had set up a business in outdoor meat BBQ. What had emerged was her experience with the 'angry vegetarian collective,' people who join a meat BBQ Facebook group or attend a BBQ event just to disrupt and demonstrate. There is a seam of anxieties, emotions and distrust running below the surface, with an observable lack of middle ground on any point of view. This didn't come up last year in what I would now characterise as earnest or polite discussions. Who else is discussing this? April Jeffries, in her PechaKucha at EPIC 2018, discussing middle-class moms in America and the pull toward insular communities. But while Jeffries' talk explored how hostile or challenging political arguments were increasingly infiltrating conversations, here were more mundane topics such as gardening, food and transport being seen as tricky and dangerous. I go back to the word 'crazy,' which has come up with different people, and remind myself how so much of our talk and our behaviours are shaped by our social circles.*

A thread through these conversations was the preference for silence in social situations outside intimate groups. Talia's dad did environmental (what she termed 'crazy' behaviours) when she was growing up (for example, cycling to work, saving shower-water for the garden). Now, as a mother in her late twenties, she is revising her opinion about her dad being 'crazy':

> Probably this weird weather, me and my partner, while we like to fish the last
> couple years it's been a little bit of a struggle. So these things directly I think
> affected me which you know is so selfish and I'm like, this is affecting me. How
> can I make change happen. Now I care Yeah, but it was like stuff like that.
> Not even being. I don't know if you're aware, but in Te Atatu. It's on the beachline.
> So we locally go out. You know, a lot of the times we've gone—we're not allowed
> to go fishing. Because of the change in like the sewage which is because of the way
> the crazy weather, so it's directly affecting me hmm and that's when I thought okay,
> maybe Dad's not a nutter you know, maybe these people are making some sense.

So then I started to ask a few more questions looked at a few more different things and thought, oh wow, this is happening, something is not right.

What is this that is happening?

…I don't know don't know if it's climate change because I still haven't come to that realisation.  There's something… I don't know…

What's the hesitation?

I guess it's…I should probably do more. If I say it's climate change I should be doing more.

Talia's realisation that something is not right derives not from any big idea or organisational intervention but from something small, local, and recent, impacting directly on her life. Yet she couldn't call it 'climate change' because she didn't know if she was 'right' and being 'wrong' is a social risk; it's easier to keep quiet. She also hesitated, because naming it 'climate change' meant that she herself would have to do more. But what could she do? What should she do? How would that impact her current way of life? Would it make enough of a difference anyway? She was overwhelmed by both the not knowingness and the scale she faced; it was easier to ignore and be silent in thought and word. We saw this throughout our work: people not talking in or with big groups about big things, and sticking to the safe topics, in our case plastics and recycling, for multiple reasons that can be summed up as 'Keeping quiet: Don't let me be the crazy one.'

People actively try to avoid being seen as crazy, fanatical or weird. These words appeared in conversation when people talked about more environmentally-minded family members, or even themselves. In casual conversation, they might phrase their position as "by no means fanatical". While they talk about their actions as meaningful, they are not discussing them in wider circles because they worry about being seen as crazy. While chatting with one participant, working in the fashion industry, about her worm farms. She told us:

I have a worm farm ...if I can contribute even in a small way, that makes me feel better. But then I guess if I, like if I went to a party or a social situation, I may not say I've been doing this and this because somebody might say, why would you do that?... Why do all these weird things?

It was important and meaningful to her, but carried enough social risk that she hesitated discussing it in a social situation. People are anxious about what people think of them, and even mentioning having a worm farm or doing something in the house to save water can label you as crazy. Elizabeth Shove (Shove 2010), in presentations on social practices, observes effort is focused on individual attitudes, behaviour, choice price and persuasion. She argues that is the dynamic regimes of everyday life; changing definitions of normal practice generate changing patterns of demand for energy, water, and other resources. Shove states cycling involves "A bicycle, a road, an ability to balance, and the sense that this is a normal and not a crazy thing to do." In conversation with us, two people dropped their voice when discussing the difficulties of mentioning these topics with wider family, let alone debating them more broadly. Only the topic of recycling seemed to make people more

comfortable. (Though, as Aucklanders, we mostly ignore the small signs of no magic recycling fairy making the waste disappear.)

It is not only fear of what people might think of them that causes people to be quiet; there is also the actual experiences people have (or the difference between expected and actual norms). This was described by Colin as a minefield—while trying to combat negative elements in a discussion, he experienced entrenching himself and other people in their position by defending his electric car purchase:

> we were lucky enough to have an electric car. So they, not they but a good portion of people, quite like, they want to tear down a new thing right? To come up with you can't tow with electric car. So well, I don't want to tow anything. Yeah, you know, I'm not saying you have to get an electric car. I'm just saying that, you know, these are the, these are the positives of an electric car. Oh, he can't go very far. And there's always negative negative, negative, negative, negative and I'm like, I try and like combat that but really all I'm doing is just entrenching myself on one side and they're entrenching on the other side.

This raises the challenge of understanding how new and different ideas enter small talk if people only talk (and listen) to the like-minded. How will new ideas about what we could be doing spread if we are retreating into safe domestic spaces and shouting into an echo-chamber? Colin raised the issue: "Do you work with people that are the same as you, or … with people that are different as you different from you so you can butt heads or you know or are you working with people that are similar so it's thinking, thinking similar." How do we break out of our echo-chambers?

## KNOWLEDGE: I DON'T KNOW THE RIGHT THING TO DO ABOUT IT

Finally, people are reluctant to talk in wider circles because they are uncertain of what is right, or as Kerrin told us, "I don't have the knowledge." This echoed our global project in 2019, where people were unsure of what the right thing to do was. They are no longer learning from verified, independent expert sources but instead from the internet. We also noticed a distrust of established mainstream media organisations who push an agenda that is 'not me.' They talked of not watching the news, and referenced Instagram or Netflix. People make sense of what is important to them, using their own sources to find information they want or need to feel like they can make an informed choice. But when they dig into topics (moving from forming opinions based on Instagram likes), they discover issues are not as straightforward as social media portrays. They don't know what the 'right' thing to do actually is; information is now a contested space. In an environment where everything is binary, there is massive social fear of sharing small talk, let alone big talk, that may result in becoming the object of a social media pile-on or 'cancelled'.

Social media might have been a space for small talk discussion, but it is increasingly tricky due to the social risk of being on the 'wrong side' in more polarised discussions. It's also a performative space. Talia mentioned how her daughter resented the family moving away from meat, but was quick to promote herself on Instagram doing "meatless Mondays." Kerrin talked of being naughty and tagging someone on Instagram who had promoted recycled coffee cups one day, and then the next day posted an image of drinking a

(throwaway) milkshake. Emma, whose husband was involved in the meat BBQ sector, talked of the herd mentality and experiencing "the angry vegetarian collective," people who join a meat BBQ Facebook group or attend a BBQ event to demonstrate. This sentiment was expressed across different scenarios through increased antagonism and a lack of middle ground on any point of view, with the entire dynamic being one of me/not me, us/not us, vegetarian good/non-vegetarian, pick your binary positions:

> I even text my husband. He's, he's quite careful because he get you know, like we've, we had a situation quite early on when we were talking about barbecue on that Facebook page. And because there's a lot of Australians on it. And someone addressing something about lamb. And my all my husband wrote was, oh, and New Zealand, lamb as you know, we love lamb and it's an inexpensive cut of meat, just made a general comment. This guy went through my husband's Facebook, got to his business page gave him a one star, wrote to him and said how do you like that arsehole. Wow. And had never met my husband and my husband hadn't said anything particularly specifically to him. Had just made a comment that we live in New Zealand, like lamb, and went against him liking lamb, and now I can't even get that removed off of Michael's business page. We've tried through Facebook and they won't remove the one. The ones in there didn't even know him. He's not even in New Zealand. You know, it's interesting the way people's minds work that they'd make that much effort just because they had a disagreement with you. And I mean, social media. I think Lady Gaga said that social media are the toilets of the internet and it's so true because there's so many people that seem to have nothing better to do with their time than to write things to other people that are just so unnecessary.

She and her husband are just trying to make a living. If, even amongst their fellow meat-eaters, stating a fact about lamb in New Zealand can hurt their business, how can you expect to talk about bigger things?

She was also genuinely puzzled why people picket at an event that's for 'meat people'. Why yes, her entire family well-being revolves around meat-based BBQs; therefore 'bad' based on the plant-based diet sustainability metric. But they also recycle the paper, plastic and food 'waste' created by these events; and she grows, consumes and bottles her own fruit and vegetables. Yet there appears to be no time, space or appetite to recognise this level of complexity. There is an increasing us versus them mentality that makes it difficult for people to talk about what they're doing and why. People are experiencing increasing polarisation— and this interferes with social survival. In contrast, organisations are increasingly pushing messages about a world where diversity is embraced, and inclusivity is the norm (in part motivated by their own social survival), a world that may only exist in marketing and communications.

This has some challenging implications. It is typical to see advice about communicating sustainability and 'keeping it simple.' But if people do not feel safe enough to engage in relatively mundane topics, how can these organisations expect people to engage in the big talk around, for example, climate change, or adapt behaviours? Increasingly, positions are held based on populist opinion/reaction, not expert/fact-based information. This contrasts to Ted Talks where we are encouraged to talk about climate change (Hayhoe 2018). But if people do not feel safe enough to engage in relatively mundane topics, how can we expect people to engage in, let alone adapt behaviours for sustainability?

This has implications for scale. One approach might be heading in a similar direction as those working successfully with co-design. They blend the small scale of teams with lived experience with larger experiments. As McKercher (2020) advocates a practice where she recommends to:

> typically limit co-design teams to 20 people to prioritise trust, intimacy and social connection. At times, you might have several co-design teams (for example, where people can't safely be in a room together or where decisions impact millions of people).In small circles, it feels easier to care for and about each other. Small groups are often less intimidating and formal than larger groups. If you're hesitant, resist falling into the trap of representation—your co-design team cannot be expected to know all things. They focus on depth, while your big circle can focus on breadth. #Tip: Resist the temptation to replace small circles with big groups, shallow consultation and one-off events.

Suchman is quoted in an article called "If You Want to Understand the Big Issues, You Need to Understand the Everyday Practices That Constitute Them" (2019). Her early research looks at how people create meaningful action by improvising:

> There is still a tendency to take those for granted, to treat those as if there is nothing to learn about the big issues by looking at mundane practices. And what does it mean to really do that. I think doing that requires a certain kind of access to the worlds that you're interested in in particular ways that can be quite demanding. To me the aspects of the book that could find their way even more actively into sociological research have to do with that commitment to the idea that if you want to understand the big issues, you need to understand the everyday practices that constitute them. (2019, 32)

These small everyday practices need to be more closely connected to the large scale efforts by sustainability practitioners inside organisations, which may require a different view of 'bang for buck' by the organisations purchasing these.

## SURVIVAL EVEN MORE IMPORTANT ON A TEMPORAL SCALE— RECONCILING OURSELVES TO A DIFFERENT FUTURE

When we talked to people about the future, they seemed to be reconciling themselves to not only an altered present, such as not being able to fish locally, but also a different future. Our final question used a different format, because we were keen to get people to actively reflect on what they told us and what it meant.

We showed people the four spaces as represented below (Hayward and Candy 2017). Hayward initially developed the grid to explore both how they saw the future (getting better or worse), and their perceptions of their own agency. We asked people, based on their conversation and their experience of the world, which quadrant best captured (or not) where they saw themselves. Did they feel things are getting better or getting worse generally; and what was their perspective of their actions?

| Upper Left: Things are good and getting better; | Upper Right: Things are good and getting better; |
|---|---|
| We have to work with the larger forces and play our part. | AND we can act to make things even better. |
| Lower Left: Things are getting worse; | Lower Right: Things are getting worse generally; |
| There is nothing I can do about it. I cannot make things worse so I am free of the responsibility of trying to do that. | But I can act to make a difference here and now, in this place. It may not change the futures but it is still worthwhile. |

© Hayward and Candy, used with permission.

We were interested to see what reaction this provoked in people about the future, and how this functioned as an alternative summing-up method. People wanted to be optimistic. Many of them would start, pause and then restart, saying, "When I first looked at this, I thought I would go with this statement 'Things are good and getting better", but as they reflected on the wider context (community, planet), they couldn't stand by it; as Talia said, "if there is genuine to homeless people living in Te Atatu how can I believe that things are getting better?." Paige reflected:

> I would probably say here. Thanks again, getting worse. But I can actually make a difference here and now. This place might not change your future but it's still worthwhile. Yeah, that's a very pessimistic, optimistic, optimistic answer. But I feel I guess I would feel worse if I didn't do anything like you know, we do little things here and there, but I would makes me feel a little wiggle a smidgen better knowing that I'm trying, you know, trying to do something rather than just doing nothing, but things that I do and they like I said, they are very small.

Paige desperately wanted to be optimistic, but her experiences said otherwise. It's quite confrontational; self-identifying as 'not a doom and gloom merchant', hoping and wanting to part of the solution, and then the realisation of the impossibility of holding this space in a visibly deteriorating world. People are becoming reconciled to a bleak future and apparently experiencing a type of grief. Sometimes labelled 'felt experience', or the 'affect theory', where people are shaped by narrative, mood or atmosphere. Or as Brown et al. (2019, page) put it, "Our emotions are a response to the way we culturally perceive the world around us." While people would like to hold onto a belief that "things are getting better", they are reaching an acknowledgement that things are getting worse.

Do not mistake this as an argument or defence for doing nothing, as this is not the case. It was interesting that our participants claimed to take actions, not because of a big or better future, but because they believed it might make a small difference for their own survival. They can, and will, do their own 'small' part even when they know it will not affect anything at a global scale. Neither the present nor the future of the planet seems okay to them. Everyone seemed to agree the future was screwed for those currently at school. Whilst younger respondents felt it more acutely, older participants were concerned future generations will face tough challenges. People grapple with legacy. This translates into the complexities of living in the now with an eye to the impact their action/inaction has on the

future for the current generation, without any certainty that their practices are the correct/right ones to make a difference.

We wonder too if there is a difference between people in design roles, or leading futures work, and our participants. Interestingly, on using this activity with different groups, Hayward and Candy (2017, 10) report that "Leaders of various kinds are often well-represented in the UR [upper right] quadrant." This is the quadrant "things are good and getting better." Other people who have used this grid as part of their futures work have commented in their blogposts that entrepreneurs also tend towards "The world is getting better and my own efforts directly affect it."

Optimism and hope are tricky. It can lead to headlines like "Engaging with the SDGs (Sustainable Development Goals can help us build a better future post-pandemic." The use of the term "better future" is intriguing. Are we in a transition? We don't want to recognise negativity but don't believe that things will get better. Is it possible if we focus on the small everyday practices, and the zones of small talk; we don't have to tightly connect actions to big changes. We asked Annie how her younger self would have answered. She told us:

> I was totally focused on myself like a 25 year old I've talked about. And if we'd asked you then would you say things getting better or things are getting worse? I would have seen probably things are getting better. Yeah. Yeah. Because I needed to believe that….Because I was terribly depressed and not in a good headspace and I had to have something to hold on to. Someone would say to me Everything is getting worse. I'm feeling well, what's the fucking point and everything's shitty. I was I would rather have had a comfortable lie than the truth. I was just like, tell me what I want to hear what I want to hear because I don't want to if it's not positive and it doesn't make me feel good and it doesn't make me feel better about my day, I'm not interested. Don't be a downer would have been my attitude you.

Turning a blind eye to what is happening is one tactic that people admitted was part of their survival repertoire.

## Have we ended up with unsustainable sustainability?

When people hear 'sustainability,' they simultaneously hear 'unsustainability'. We saw that above with Talia and work. At the end of our conversation with Annie, we were interested that the word hadn't come up given what we had talked about. So, as we were finishing, we asked about the word 'sustainability'. She frowned, let out a sigh:

> I guess in the past, people have talked about. They've got these big ideas about and this may be completely wrong. And gosh, I'm really scrambling.
>
> You said you haven't heard in a while?
>
> Yeah, I thought this has been so off my radar with what's been going on the last few years ...sustainability. Probably a few years ago, people were talking about how to despite the environmental demise and the way the world's headed, finding ways to keep crops going or keep supply and demand running along well, and you had this big shift in veganism where everyone's Let's all go plant based and let's make vegan meat and so people are either wanting to keep bleeding everything dry

forever, or they going let's find alternative sources now. That's about the only link I have to sustainability that comes into my mind. I haven't heard it for a while….

So any of the stuff that you do for example, when you say you're trying to do your bit, reducing your use of plastics no longer having a bin liner doesn't fit under sustainability for you?

No, not really.

Psychologists use the term 'semantic satiation' for when the repeated use of a word loses any meaning. Like wallpaper, we see it every day but are no longer seeing it.

## SURVIVABILITY IS MORE MEANINGFUL

When we started this project, we were interested in what language people might be using. What we didn't expect to hear as a cross-current through the project was the concept of survivability. We ourselves submitted the initial abstract for this paper with the statement, "Sustainability is the largest scale issue for the planet. Sustainability is a global challenge. It's important to us as human beings and ethnographers." When we interrogated this, we realised that not only was 'sustainability' not homogeneous to our participants, but also that it did not represent the things that people do to make themselves feel less bad. People don't use the term 'sustain' or raise the question of who gets to sustain what. The conversation was actually about survivability, not sustainability. Participants used 'survive' as a verb in contrast to 'sustainability' as a vague noun. Surviving is much more tangible and includes getting through life as best as you can. Not just the wherewithal to provide you and yours with shelter, food, water; but also the social intelligence, tools and dexterity to live risk-free in wider society. People had their own ways of coping with what is happening with the planet, including ignoring it. Survival was also framed within the household: "if you're a family struggling you know to make ends meet, then you're going to buy what you can."

We looked around to see if other people were discussing sustainability and survivability. In the introduction to the *Extinction Rebellion Handbook*, Knights writes, "we acknowledge that Extinction Rebellion is just one articulation of a feeling that is being felt all across the world… To survive it's going to take everything we've got" (2019, 12). Peña-Taylor, on the World Advertising Research Centre website, headlined an article with 'Sustainability is about survival—the subject needs to overcome the semantic bleaching that makes it a corporate nice-to-have." He says, "Like the other instances of semantic bleaching that the corporate world is so fantastic at perpetuating (think how 'undoing structural racism and sexism' became 'diversity'), sustainability needs to be more associated with survival than anything cuddly." Returning to anthropological literature, we came across a chapter on sustainability in the book *Lexicon for an Anthropocene Yet Unseen*. It includes a section where a team is trying to learn about the social lives of sustainability (as if sustainability had social lives) in Guatemala and grappling with translating the word sustainability:

We worked together with Marta on the term sustainability for a while—co-laboring, to use Marisol de la Cadena's (2015) term for a collective effort to attend to spaces of difference. We slowed down when we came across tanquib'ela, which back-translated into el ser en la vida; de vivir; de sobrevivencia, and then, being in

life, of living, of survival...The shift between tanquib'ela, mantenerse, surviving, or improving the world is not innocent, but forecloses some worlds while attending to others. The challenge we faced was not only a matter of moving between English, Spanish, or Mam. Whereas a lexicon may elsewhere be units of meanings, the anthropologist's lexicon might be better conceived as a repertoire of care-filled practices that follow the conversion of spoken concepts into unspoken activities and back into words again. Engaging in the practice of the lexicon requires the skill of asking—and sometimes not asking (see Pigg 2013)—about how words are done and what they then do. (Meza and Yates-Doerr 2020, 465)

We realised that the question of sustainability is in part a linguistic challenge involving the role of words. We were intrigued to discover that "The shift between tanquib'ela, mantenerse, surviving, or improving the world is not innocent, but forecloses some worlds while attending to others" (Meza and Yates-Doerr 2020, 465). In Spanish, and in Guatemala, a team separately arrived at similar discussions. There is a similar level of guilt, sense of loss and acceptance. The passage above highlights "care-filled practices" versus sustainability discourse implying solutions. How might we acknowledge survival, and how does this plays out in a context of small talk?

## RECONCILING OURSELVES TO A DIFFERENT FUTURE

### Field Vignette: 3

*We are trying to rewrite a section of our paper for the umpteenth time. Louisa suddenly says to me, 'It's been two years we've been working on this and we still can't say what sustainability means.' I reply, 'Your point being?' We laugh so hard tears appear. We do not know what makes it all so difficult—'sustainability' or writing a paper during a global pandemic or something else? We do know that despite decades of global sustainability discourse the world is still going to hell. What's gone wrong?*

*We talk about wilful ignorance by us as individuals, us as researchers and us as people working with global organisations. We could stand up and deliver a pretty good talk on sustainability as business as usual. We could easily say brands need to be more concrete in how they relate to community, people are worried about the plastic in the oceans, they feel good about plastic, they say they have the seen the ads with the biggest media spend, and this is what doing good or the future looks like. It's conventional big talk. This misses what we have learned through the small talk. What follows is some tentative stuff, and questions we are left with…*

After two years of large and small projects on sustainability, what have we learned? What would we tell people we work with who are keen to know not only the 'what', but also the 'so what' and 'now what'?

### Our Own Unsustainable Use of Language

The initial challenge we discovered with sustainability was that the concept lacked clarity. There were so many different meanings associated with the word it became essentially meaningless. We also learned that the word is overused. People not responsible for sustainability as part of their role in an organisation or group are cynical about its use due to what they see on social media and their experience with organisations. Further, the use of 'sustainability' is possibly as unsustainable as the way we are living—it is not getting us closer

to providing a solution to the problems we face. Which raises a catalytic question: why do we ignore this? Why are we complicit in maintaining a social silence?

## Currently 'Sustainability' Acts As A Cipher

Sustainability appears to operate as a code word, and that social silences continue to exist inside organisations. Gillian Tett discussed such "social silences" in *The silo effect: why putting everything in its place isn't such a bright idea* (Tett 2015, 45):

> what really matters in a society's mental map is not simply what is publicly and overtly stated, but what is not discussed. Social silences matter. The system ends up being propped up because it seems natural to leave certain topics ignored, since these issues have become labeled as dull, taboo, obvious, or impolite. …as Bourdieu said: "The most powerful forms of ideological effect are those which need no words, but merely a complicitous silence.

Tett has described the social silence of people operating within the wider financial field around the explosion of derivatives in the financial markets, which led in part to the financial crisis back in 2009. We wondered: is sustainability a social silence? Why do organisations continue to communicate sustainability, even while knowing that the word is problematic? If people inside organisations name what an organisation does as 'sustainability,' this is a form of power or as Bourdieu would term it, a use of symbolic capital. This use of language within organisations leads to Annie and Talia being disillusioned with 'sustainability' as they notice that its primary use is not for them but for the users of that language to claim a particular status.

Practically, sustainability appears to now operate as a cipher—you only access value by knowing how to use the code. The word 'cipher' has an Arabic root. Sifr; zero, empty, nothing. When a word like 'sustainability' is used so broadly, it becomes an empty signifier. Holly Jean Buck – Assistant Professor of Environment and Sustainability University at Buffalo in New York – discussed this when participating in an online conversation at Assembly 2020. She was invited along with others because her specialist work embodies long term thought to discuss the future. When asked about the political ways in which a word is used, she was specifically asked:

> in the vernacular of your field, what words, what phrases what terms have become bastardised, have become overcomplicated oversimplified so they have no meaning and now their meanings have been galloped away with?

She replied: "Sustainability- it's become empty of content….it's become a corporate term, it's a work phrase now"( 2020, 8:10:12). The in-group is less concerned that the word lacks concrete meaning, as it performs its function as code for funding, measures or proof of corporate citizenry etc. Restricted codes have always been used. Sentries and doorkeepers may only allow someone to pass if they know the code. Thus, codes operate as secret signals or 'badges of belonging'. Sustainability and its derivatives sometimes operate as such a code where you need to decipher intention and appropriate behaviours. No longer do people need a special word to gain physical entry; they need the appropriate language to be considered a legitimate 'player' for funding, recognition, or inclusion for their own survival. Hence we see the use of vague phrases such as 'business sustainability review,' or 'sustainable economic

growth.' Others see the word, with limited understanding or appreciation of its intended meaning or purpose, and use it to say: "I'm a part of this club too." The word loses legitimacy or potency through inappropriate use.

How about a different word? As a candidate, how about 'survivability' as a more representative word capturing how real people are thinking, talking and acting? We are hard-wired to survive. Looking across the vast temporal scale of human existence, since our ancestors first crawled out of the sea, we humans have continually adopted and adapted the actions and behaviours required to survive. Let's not forget that organisations are driven to survive also, taking on board the policies, practices and publicity that keeps them viable. It's survival of the fittest, not sustainability of the fittest.

## If You're Trying To Talk To People About Sustainability, You Can't Talk About Sustainability

The 2030 Agenda for Sustainable Development was adopted in the UN General Assembly in 2015 and has primarily been picked up by governments and businesses, either for action or justification. In May 2018, during a UNEP/OECD event on sustainable lifestyles, Townsend, of Futerra, developed a version of the UN Sustainable Development Goals to make them more tangible. An example is instead of "reduced inequalities" she used "Be fair, " instead of "Sustainable Cities and communities" she used "Love where you live." They were intended to be clearer and more engaging than the original goals and proved popular when she shared them on social media. After some copyright issues following a tweet, there was a collaboration between Futerra and the UN resulting in "The Good Life Goals", which were launched at the UN headquarters in 2018 focused on "bridging the gap". What is the gap? Futerra characterises this gap on their own website as a language gap. "The world's governments, civil society, stakeholders and business did a pretty good job of creating a To Do List for humanity. Except that list is written in a way which excludes the most important change-maker of all—YOU." The Good Life Goals do not mention sustainability and are characterised by concrete and specific language. Examples of the rewritten goals are below: (Futerra, 2020)

The criteria for rewriting the goals included:
1. Will this personal action make a tangible **impact** on the Goal?
2. Will this action be **accessible/relevant/affordable** to the greatest number of people?
3. Is the action **comprehensible** and likely to benefit our own lives? (Futerra, 2018)

The reworked goals are much closer to the social practices of our participants – small, attainable, tangible and socially acceptable rather than increasing their risk of being a target at the next social gathering. Further, they are an example of what Townsend believes that "*I personally believe that people power is as important as powerful people.*" She and her team are actively bridging the gap between people and very large organisations, such as the United Nations.

**Holding the Space Between an Optimistic and Pessimistic View of the Future**

People inside organisations tend to use 'sustainability' in an optimistic way. As described by Meza and Yates-Doerr:

> Sustainability (and, we might add, becoming and emerging, since these terms often go hand-in-hand) may too easily connote the progressive transition of a singular, causal system, leading us toward the project of developing a better future that has long been modernity's destructive lure (2020, 465).

Hayward's phrase that we used "Things are getting worse generally… but I can act to make a difference here and now, in this place" is close to what people are telling us about their experience. "It may not change the future, but it is still worthwhile." It turns out people don't have to feel that they can make a difference to the world to act in the here and now. Remember that our participants claimed to take actions, not because of a big or better future, but because they believed it might make a small difference for their own survival. It might help their physical survival, probably their social survival, and at the least will make them feel better about themselves. They can, and will, do their own 'small' part even when they know it will not affect anything at a global scale. How do we design new products and services and communicate these in a world where people want to be optimistic, but can't? If we accept ethnographers have always been engaged in the language of loss (Behar 2003), then some of the answers and questions will lie in our own practice.

## EMBRACING THE SMALL TALK ABOUT SMALL THINGS

Bronislaw Malinowski introduced the term 'phatic communion', verbal or non-verbal communication that has a social function, nearly a century ago in his essay "The Problem of Meaning in Primitive Languages" (1923). We found that small talk is more than just chit chat; it's a way of navigating massive issues. Start where people are at; small and concrete. Our brains struggle to hold the complexity of the large scale; we naturally gravitate to the small scale of the local.

We notice how much emotion and subjectivity are caught up in people's reactions and behaviours. We contend that, in some pivotal way, sustainability comes between us and our relationship with our environments. The OED identifies the primary meaning of sustainability, and the verb 'sustain' as 'to keep in existence, maintain; spec. to cause to continue in a certain state for an extended period or without interruption." Our participants know, without being able to articulate this, that the current way of living is unable in some way to continue. Survivability has more in common with an obsolete form of 'sustain" which derives from Old French sostenir, meaning "hold up bear; suffer, endure."

While the people we spoke to are adjusting to the signals of a darker future, they themselves see both meaning and hope in both their small talk and their actions. Small talk may be about intimate and small scale things, but in fact, it is important on a human level and is probably also pretty widespread. There's a political choice not to use what has become bureaucratic, abstract big talk, but rather to lean closer to both worlds and the conversations that people are having in their own local situations. We wonder about the implications for

action. We have learned we need to embrace the mundanity of the small talk if we don't want to sabotage our survival.

**Louisa Wood** is happiest immersed in unravelling the human experiences of others. Her foundation in theatre and semiotic thinking provides a different attentiveness to the framing of these, while a career in qualitative research provides a space in which to do so.

**Lee Ryan** completed an MA Hons at the University of Auckland then another at UCL. She was the Regional Director for Qualitative Research at TNS for Asia Pacific, Latin America, Middle East & Africa. Back in Aotearoa, she aims to write more things long-form after writing 4125 PowerPoint slides.

## NOTES

## REFERENCES CITED

Auckland, Emily. 2020. "Engaging with the SDGs can help us build a better future post-pandemic." *Edie* website, May 4. Accessed [September 1, 2020]
https://www.edie.net/blog/Engaging-with-the-SDGs-can-help-us-build-a-better-future-post-pandemic/6098765

Behar, Ruth. 2003. "Ethnography and the Book that was Lost.", *Ethnography* 4 (1): 15-39.

Bourdieu, Pierre. 1994b. *The Field of Cultural Production: Essays on Art and Literature.* New York: Columbia University Press.

Brown, Steven D. Kanyeredzi, Ava, McGrath, Laura, and Ian Tucker. 2019. "Affect theory and the concept of atmosphere." *Distinktion* 20 (1): 5–24.

Buck, Holly Jean. 2020. "Longplayer Assembly." *Artangel* website, 2020. Accessed [September 27, 2020].
https://www.artangel.org.uk/longplayer/longplayer-assembly/8:10:12

Colmar Brunton. 2019. "Better Futures: Celebrating a decade of tracking New Zealanders' attitudes & behaviours around sustainability." *Colmar Brunton* website, 2019. Accessed [September 7, 2020].
https://static.colmarbrunton.co.nz/wp-content/uploads/2019/05/Colmar-Brunton-Better-Futures-2019-MASTER-FINAL-REPORT.pdf

Futerra. n.d. "Why We Need New 'Good Life' Goals." *Futerra* website, 2020. Accessed [September 1, 2020].
https://www.wearefuterra.com/2018/07/why-we-need-new-good-life-goals/

Galik, Gyorgyi. 2019. "The Change Before Behaviour: Closing the value-action gap using a digital social companion." *Ethnographic Praxis in Industry Conference Proceedings,* 2019: 415-439.

Ghosh, Iman. 2020. "Visualizing the Global Rise of Sustainable Investing." *Visual Capitalist* website, February 4. Accessed [September 1, 2020].
https://www.visualcapitalist.com/rise-of-sustainable-investing/

Goffman, Erving. 1959. *The Presentation of Self in Everyday Life*. New York: Anchor Books.

Hasbrouck Jay, and Charlie Scull. 2014. "Hook to Plate Social Entrepreneurship: An Ethnographic Approach to Driving Sustainable Change in the Global Fishing Industry." In: *Handbook of Anthropology in Business, edited by* Rita Denny and Patricia Sunderland, 186-201. Walnut Creek, CA: Left Coast Press.

Hayhoe, Katharine. 2018. "The most important thing you can do to fight climate change: talk about it." *TED* website, November 2018. Accessed [September 1, 2020].
https://www.ted.com/talks/katharine_hayhoe_the_most_important_thing_you_can_do_to_fight_climate_change_talk_about_it

Hayward, Peter, and Stuart Candy. 2017. "The Polak Game, or: Where do you stand?". *Journal of Futures Studies*, 22 (2): 5-14.

Hunter, Laura. n.d. "Laura Hunter." *Intent* website, 2020. Accessed [September 30, 2020]
https://www.intentjournal.com/features/laura-hunter/

Inagaki, Kana. 2020. "UN sustainability goals set standard for Japanese businesses." *Financial Times* website, July 23, 2020. Accessed [September 1, 2020].
https://www.ft.com/content/89711e3c-afeb-11ea-94fc-9a676a727e5a

Kaplan, Josh. 2017. "When the Emperor Has No Clothes: Performance, Complicity and Constraints on Communication in Corporate Attempts at Innovation." *Ethnographic Praxis in Industry Conference Proceedings,* 2017: 217-231.

Kimbell, Lucy. 2014. "Design Ethnography, Public Policy, & Public Services: Rendering Collective Issues Doable & at Human Scale." In *Handbook of Anthropology in Business*, edited by Rita Denny and Patricia Sunderland, 186-201. Walnut Creek, CA: Left Coast Press.

Maldonado García, María, García Meza, Rosario, and Emily Yates-Doerr. 2020. "Sustainability." In *Anthropocene Unseen: A Lexicon*, edited by Howe Cymene and Pandian Anand, 465-469. Earth, Milky Way: Punctum Books.

Malinowski, Bronislaw (1923) The problem of meaning in primitive languages. In C.K. Ogden & A.I. Richards, The Meaning of Meaning, Supplement I. London: Kegan Paul, pp. 451-510.

McKercher, Kelly Ann. 2020. *Beyond Sticky Notes: Doing Co-design for Real: Mindsets, methods and movements*. Sydney, Australia: Beyond Sticky Notes.

McGonigal, Jane and Mark Frauenfelder. 2016. "Futurist Imagination Retreat Report." *University of Pennsylvania* website, May, 2016. Accessed [September 1, 2020]
https://cpb-us-w2.wpmucdn.com/web.sas.upenn.edu/dist/8/471/files/2018/06/Futurist_Imagination_Retreat_Report-2fge730.pdf

Moore, Henrietta L. 2004. "Global anxieties: concept-metaphors and pre-theoretical commitments in anthropology." *Anthropological Theory* 4 (1): 71-88.

Nielsen. 2018. "The education of the sustainable mindset." *Nielsen* website, 2018. Accessed [7 September 2020].
https://www.nielsen.com/us/en/insights/report/2018/the-education-of-the-sustainable-mindset/

Oatley. 2018. "Sustainability Report 2017." *Oatly* website, October. Accessed [September 4, 2020]
https://www.oatly.com/uploads/attachments/cjp9gpbd709c3mnqr8idyecwy-sustainability-report-2017-eng.pdf

Shove, Elizabeth. 2010. "Beyond the ABC: How social science can help climate change policy." *SlidePlayer* website, n.d. Accessed [September 27, 2020].
https://slideplayer.com/slide/1658232/

Suchman, Lucy, Gerst, Dominik, and Hannes Krämer. 2019. "'If you want to understand the big issues, you need to understand the everyday practices that constitute them.' Lucy Suchman in conversation with Dominik Gerst & Hannes Krämer." *Forum Qualitative Sozialforschung / Forum: Qualitative Social Research* [Online] 20 (2). Accessed [September 1, 2020].
http://dx.doi.org/10.17169/fqs-20.2.3252

Tett, Gillian. 2015. *The silo effect: why putting everything in its place isn't such a bright idea.* London, U.K.: Little, Brown.

Townsend, Solitaire. 2018. "An approach to making the SDGs personal." *GreenBiz* website, October 4. Accessed [September 1, 2020].
https://www.greenbiz.com/article/approach-making-sdgs-personal

Toyota. 2020. "Sustainability." *Toyota* website, 2020. Accessed [September 4, 2020].
https://global.toyota/en/sustainability/

Tsing, Anna, 2017. "A Threat to Holocene Resurgence is a Threat to Livability." In *The Anthropology of Sustainability: Beyond Development and Progress*, edited by M. Brightman and J. Lewis, 51-66. Volume in *Palgrave Studies in the Anthropology of Sustainability.* New York: Palgrave.

Van der Linden, Sander, Maibach, Edward and Anthony Leiserowitz. 2015 "How to Improve Public Engagement with Climate Change: Five "Best Practice" Insights from Psychological Science". *Perspectives on Psychological Science* [Online] 10 (6): 758-763. Accessed [September 7, 2020].
https://scholar.princeton.edu/sites/default/files/slinden/files/ppsfinal.pdf

Youngblood, Mike. 2016. "Sustainability and Ethnography in Business: Identifying Opportunity in Troubled Times." *Epic* website, December 5. Accessed [September 4, 2020].
https://www.epicpeople.org/sustainability-ethnography/

**THEMATIC SESSION**

# Isolation and Connection

This session draws on experiences ranging from early ethnographic accounts of life on islands, COVID-19 hospital wards, and the International Space Station to explore how we work and live in a period of enforced distancing and isolation.

Session Curators: Evan Hanover, Thomas Lee, Elena O'Curry, Tabitha Steager

# Postcards from Isolation

## Digital Artefacts from the Lockdown Time

ANNA WOJNAROWSKA, *Google*

*'Postcards from Isolation' is a collection of collaborative & interactive experiences that represent the shifts caused by the COVID-19 pandemic.*

*The origins of the project lie in a reflection on the deeper social, cultural and economical impact of being in isolation. Drawing on ethnographic accounts (Malinowski, Mead), we studied how the context of being isolated these days corresponds to the uniqueness of geographical islands - spaces that enforce a different connection to land and its resources; and ones that lead to unique cultural patterns. Drawing on that metaphor, we explored to what extent we, as individuals, may be evolving into symbolic islands of our own, driven by similar powers, at a scale unlike before. What can we learn from that? How can we visualise these shifts for others to remember?*

*The non profit project was self-initiated by a group of cross disciplinary HCD colleagues. The website (isolation.is) currently holds 15 Postcards that represent shifts most important to its contributors. This PechaKucha tells a story of how Postcards became (1) a platform that allows for a rich understanding of how the COVID pandemic is shifting our lives; (2) a living repository of digital artefacts that can be studied like any other piece of material culture; and (3) a collection of digital drawings that are an important asset in any anthropologists' notebook (Taussig).*

*Keywords: digital artefacts, isolation, islands*



Home page of Postcards from Isolation. All experiments can be found at https://isolation.is.

**Anna Wojnarowska** is a senior user researcher, currently working at Google. She is informing the design of health and wellbeing products by applying ethnographic methods to global research projects. Before Google she worked at a design consultancy, Experientia, and spent years freelancing in the public sector in the UK - with central government, local councils and the NHS being some of her previous clients. Recently she started applying her research skills and expertise onto conceptual side projects like Postcards.

# Architecture Can Heal

## Spatial Literacy to Protect COVID-19 Healthcare Workers

MICHAEL DOLINGER, MD, MBA, *Mount Sinai Kravis Children's Hospital*
ASHLEY MARSH, RA, *MASS Design Group*

*In April 2020, a study of The Mount Sinai Hospital in New York City was conducted to better understand the challenge of adapting idealized infection control design guides to site-specific conditions during a pandemic. The study aimed to capture quick interventions that are working, offer a new hypothesis and framework to guide future design interventions, and share lessons to assist other medical facilities as they pursue their own necessary spatial adaptations moving forward. Three units repurposed for COVID-19 were studied. Using action cameras and cloud-based videoconferencing, clinicians helped designers remotely peer in real time to active COVID-19 units, create "heatmap" annotations of perceived risk by frontline clinicians, and conduct interviews with decision makers.*

*The COVID-19 pandemic has challenged health care systems around the world to provide safe and effective care. Leveraging spatial design, architecture, and design hacks offers an untapped opportunity to support infection prevention and improved team dynamics, ultimately improving the safety and the effectiveness of the health care team by creating an environment that supports infection prevention and team function.*

## THE CHALLENGE

The novel coronavirus (COVID-19) has posed an unprecedented challenge to healthcare infrastructure around the world. What Dr. Paul Farmer calls "staff, stuff, space, and systems" have been pushed to their limit. There have been many stories about the crisis of stuff, staff, and systems, but fewer discussions about space. One reality that links them together is that healthcare workers—who are particularly vulnerable to exposure and more severe infection —must remain healthy in order to protect patients, families, and communities. And the design of hospital spaces must help protect us.

The space of the hospital itself however, will continue to facilitate nosocomial (hospital borne) infection, unless infection control protocols are established and adhered to (Liu 2020). And herein lies the problem: validated protocols designed to prevent COVID-19 transmission do not yet exist, and therefore hospitals are implementing spatial redesigns on the fly, doing their best to learn from protocols based on other diseases, like Ebola and tuberculosis.

Until we can better understand the virus' pathways, we won't be able to confidently redesign our existing spaces to adhere to new and higher standards of infection control guidelines. In the meantime, hospitals will continue to repurpose and convert their spaces adhoc to meet surge demand—adapting idealized infection control protocols to unideal spaces and situations. This means healthcare workers and administrators must quickly adapt inflexible spaces, recognizing that the resulting adaptations may put healthcare workers and

our communities at risk unless we can quickly create site-specific guidelines that are adherable and implementable based on the best available knowledge. While we need research to understand who is at increased risk for complications of COVID-19 and to develop effective vaccines and best therapies, we also need research that identifies how spatial design and awareness can mitigate risk.

## LEARNING FROM A FRONTLINE HOSPITAL

The following case is an example of how clinicians at The Mount Sinai Hospital in New York City, researchers, and designers came together to understand the constraints of adapting guidelines to imperfect spaces, and the lessons we believe other institutions can learn from this. Shortly after the Shelter-in-Place order was issued by the Governor of New York in mid-March, Mount Sinai transformed its entire care delivery system and carried out significant infrastructure modifications to address the COVID-19 outbreak.



Figure 1: Mount Sinai Kravis Children's Hospital 4th Floor South Wing on April 6, 2020, photograph copyright Michael Dolinger.

This case study presents a three-week spatial research methodology that took place within a frontline hospital at the epicenter of the COVID pandemic in New York City, with the goals of:

bringing light to the challenge of adapting idealized design guides to site-specific conditions,

capturing quick interventions that are working,

offering a new hypothesis and framework to guide future design interventions, and

sharing our lessons to assist other medical facilities as they pursue their own necessary spatial adaptations moving forward.

**METHODOLOGY**

During the week of April 6, 2020, a team of clinicians and designers from the Mount Sinai Kravis Children's Hospital and The Mount Sinai Hospital, MASS Design Group, and Ariadne Labs investigated spaces converted and repurposed for the care of critically ill COVID-19 patients. Drawing from MASS's infection control design experience and evidence-based research experience, Ariadne Lab's public health innovation pathway and knowledge sharing model, and Mount Sinai's position on the frontlines of the NYC epidemic, the team sought to illuminate COVID-19 capacity planning, interventions, and opportunities.
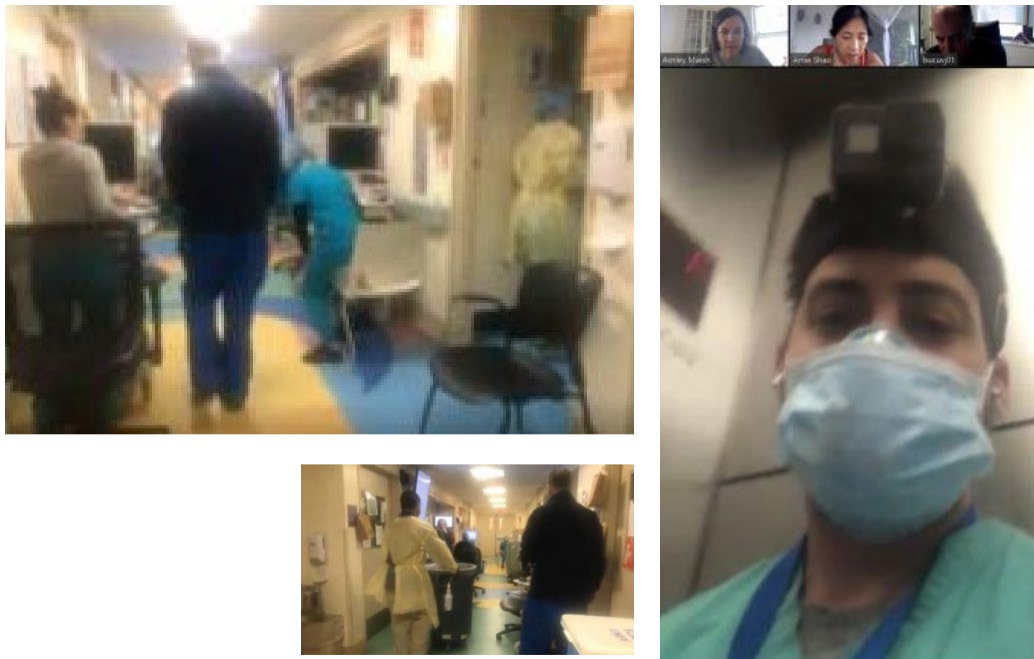


Figure 2: Screenshots from virtual walkthrough of active COVID-19 patient care unit, images copyright MASS Design Group.

The rapid response spatial study focused on three units converted and repurposed for COVID-19, including adult intensive care units (ICU), an adult medical-surgical unit, and a pediatric unit. Using action cameras (GoPro) and cloud-based videoconferencing (Zoom), Mount Sinai clinicians helped designers remotely peer into active COVID-19 units, create "heatmap" annotations of perceived risk by frontline clinicians, and conduct interviews with Clinical Operations and Facilities leadership.

Figure 3: The team documented observations through on-the-fly sketches, drawing exercises, and photographs; and also generated graphics like the diagram to the far right to better depict conditions and recommendations, images copyright MASS Design Group.

Together, the practitioners, designers, and researchers identified and documented interventions, discussed our understanding of limitations, and recommended next steps to furthering a shared understanding of design interventions and strategies.

## PRE-COVID-19 DESIGN STANDARDS

Many design recommendations from before the current coronavirus pandemic exist regarding idealized conditions for treating patients and protecting healthcare workers.

For highly contagious infectious diseases, negative pressure isolation rooms are the standard of care. These are single patient rooms designed with antechambers and specialized airflow precautions. To prevent contaminated air from spreading into the rest of the hospital, rooms must be fully sealed, achieve a specific number of air changes per hour, ensure negative pressure (preventing room air from entering the hallway), and safely exhaust contaminated air. For diseases spread through direct contact, like Ebola, other infection control precautions are layered on—such as carefully choreographed routes for patient and staff movement across the facility and special surface sanitization measures.

Until recently, however, hospitals have accommodated infectious disease patients as the exception, not the norm. Most large urban hospitals in America have a handful of negative pressure isolation rooms. Prior to COVID-19, and during the very early stages of the outbreak, this was enough. But what happens when a hospital needs to be able to accommodate hundreds of patients under pandemic conditions? This is what Mount Sinai faced in March.
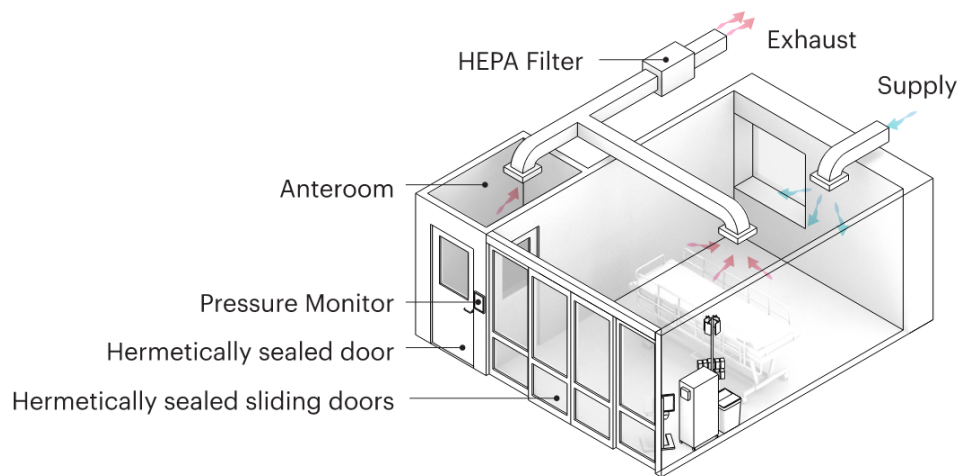
Figure 4: Diagram representing the setup of a negative pressure isolation room, image copyright MASS Design Group.

## THE MOUNT SINAI HOSPITAL

In less than three weeks, beginning in mid-March, the Mount Sinai leadership redeployed healthcare workers to COVID-19 units, backfilled the gaps by shifting provider responsibilities, implemented a new shift system to minimize risk to health workers during shift change, established complex PPE distribution and recycling systems, implemented donning/doffing practices, postponed elective procedures to create capacity for the coronavirus surge, rapidly implemented telehealth services, restructured research efforts to focus on the pandemic, and committed resources to provide ongoing support for frontline workers including childcare, transportation, and psycho-social support (Edmondson 2017).

During the same period of time, the Infection Prevention and Clinical Operations and Facilities teams at The Mount Sinai Hospital worked to significantly expand the hospital's COVID-19 bed capacity in anticipation of a surge. The novel coronavirus forced healthcare facilities to broaden their focus from the individual patient room to the entire care unit. Mount Sinai built a 100-bed step-down care unit in the hospital's atrium (Figure 5), and in partnership with Samaritan's Purse erected a temporary 68-bed tent facility (Figure 6).

Architecture Can Heal – Dolinger & Marsh

Figure 5: Photo showing the step-down care unit, photograph copyright Michael Dolinger.



Figure 6: Photo showing the temporary tent facility, photograph copyright Michael Dolinger.

## SPATIAL MODIFICATIONS DURING THE SURGE

The hospital additionally converted 260 existing patient rooms into negative pressure isolation rooms. Within adult ICUs designated for COVID-19 care, tremendous changes took place. Walls and doors were constructed across the front of previously open ICU bays, and measures were taken to ensure fully sealed patient spaces. Large high-efficiency particulate air (HEPA) filter units were brought into rooms and set up to directly exhaust out windows that were fitted with wood panels. Intravenous fluids and patient monitors were pulled out into hallways to minimize provider exposure and allow for safer and easier monitoring (Figure 7). PPE waste and recycling bins were dedicated to each unit, and protocols were established to bring these materials to dedicated service elevators in an effort to minimize cross-contamination.

Figure 7: IV fluids and monitors in the hallways, photograph copyright Michael Dolinger.



*Before: 6 isolation rooms*     *After: 48 isolation rooms*

Negative pressure isolation rooms

Figure 8: Negative pressure isolation rooms on 6th floor of Guggenheim Pavilion, image copyright MASS Design Group.
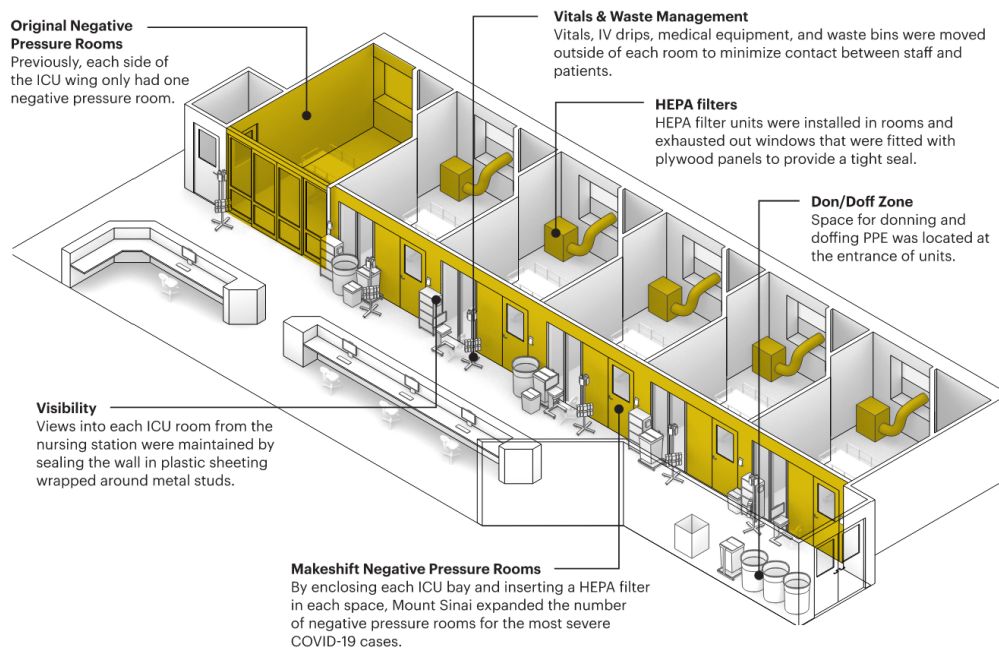
Figure 9: Diagram of interventions on the Guggenheim Pavilion highlighting different interventions that were made to turn this adult ICU wing into one that could treat patients who have tested positive for the coronavirus, image copyright MASS Design Group.

## DESIGN HACKS

Mount Sinai's administrative, Clinical Operations, and Infection Prevention teams worked together to create physical changes to reinforce the programmatic and behavioral protocols implemented on COVID-19 floors. Specifically, they sought to include visual cues at key moments of risk or transition: for example, doors into patient rooms, PPE storage and access locations, and thresholds between different areas of risk. In some places, they installed signage indicating "extended use PPE zones," designated boxes for recycling face masks, and installed additional hand sanitizers to signal to staff upon entering or exiting key thresholds. These physical or cognitive "anchors" helped to designate functional use of space and reinforce patterns of work within these constraints. Other anchoring mechanisms could include tape on the floor, signage, and intentionally placed whiteboards or paint.

Stations to access PPE were provided at key locations, which in turn affected how and where healthcare providers utilized the equipment. PPE carts and instructional signage were located on every COVID-19 unit, a brown bag system was created to designate and store face masks for individual providers, and a peer approval system was implemented—requiring nurses to fill out a donning/doffing record sheet with "buddies" signing off on proper PPE application (Zimring 2018). Additional interventions that may increase PPE adherence in the future include: demarcating designated "clean" zones around PPE carts and donning areas, creating more predictable placements for essential equipment and resources, making access to PPE less high-touch, and developing a more consistent system for brown bag storage (either consolidated near the unit entrance, or affiliated with specific patient rooms).

**Entry into patient rooms**

Hand Sanitizer
Paint indicating risk thresholds
Information Board
Tape on floor

**Exit from patient care unit**

Hand Sanitizer
Paint indicating risk thresholds
Waste Disposal
Tape on floor

**PPE access & storage**

Paint indicating risk thresholds
PPE
Hooks for hanging PPE brown bags
Instructional signage
Low-touch drawers
Tape on floor

**Objectives**
- Create awareness around what risk zones staff are entering vs exiting
- Convey key information at a glance (such as whether patients are intubated)
- Increase staff efficiency

**Objectives**
- Create awareness around what risk zone staff are exiting vs entering
- Remind staff to get clean before leaving contaminated zones
- Reduce unnecessary movement between units
- Reduce disease transmission to non-COVID units

**Objectives**
- Create clear clean zones around PPE carts and donning areas
- Consolidate needed resources in one place
- Increase PPE adherence
- Reduce time and hassle in donning PPE
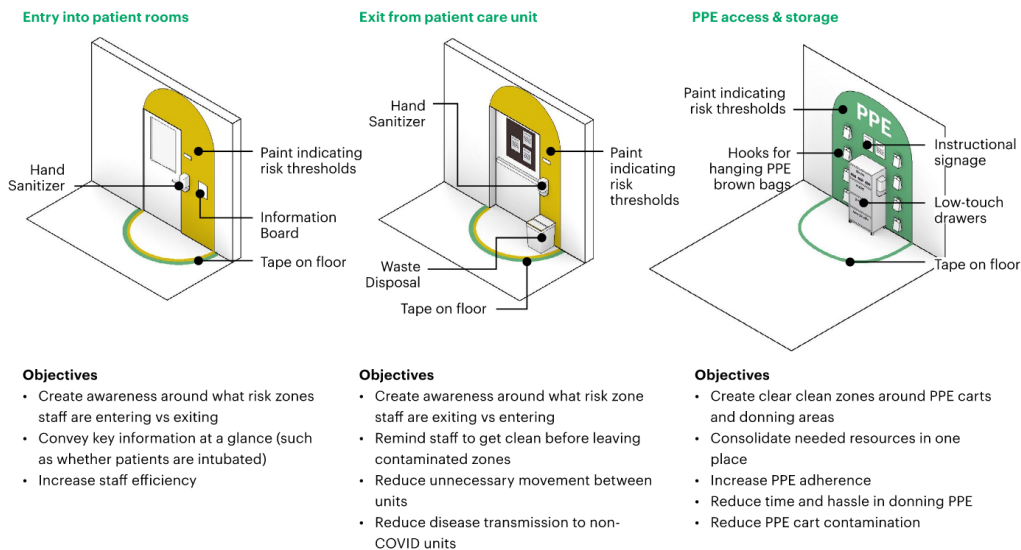- Reduce PPE cart contamination

Figure 10: Design Hacks, images copyright MASS Design Group.

## COMPLIANCE GAPS AND IMPLEMENTATION CHALLENGES

Even with Mount Sinai's ingenuity and innovation in achieving a radical transformation in a short timeframe, adapting the existing hospital proved to be a difficult task. Changes were ultimately bound by inherited and inflexible infrastructure. Within adult ICU units, it was not possible to integrate antechambers due to the spatial constraints of the floorplan, as well as the need to maintain direct visibility and access to patient rooms for high-level nursing care. Shifting IVs and monitors outside the patient room was necessary, but created more congestion in the corridors. And while the HEPA filtration fixtures were necessary, they took up an immense amount of space within patient rooms and were loud, constraining staff movement, communication, and workflows.

In older, smaller buildings like those that housed the pediatric units, repurposing spaces for coronavirus care was even more challenging. The presence of larger clinical care teams (reflecting higher nurse-to-patient ratios and more ancillary staff), alongside the need to minimize staff exposure within patient rooms, meant that more providers were working in hallways and compromising spatial distancing.

In buildings laid out without centralized nursing stations, this prompted the distribution of individual nurse computer stations along the hallway. Combined with the addition of PPE carts and recycling bins, as well as more complex cleaning and disposal protocols, the result was hallways crowded with equipment (Figure 11). Particularly in critically ill patient spaces, this left little room for clinician rounding to comply with social distancing recommendations, and introduced inefficiencies that made nurses' work slower.

Figure 11: Mount Sinai Kravis Children's Hospital 4th Floor South Wing, photograph copyright Michael Dolinger.

Beyond the size, layout, and flexibility of units, spatial perception emerged as another important factor. Clinician discussions and floor plan annotation exercises revealed inconsistent perceptions of risk zones within units, and the need for stronger spatial cues. Unlike typical care contexts, where providers work on units they are familiar with, and with teams they are familiar with, COVID-19 providers may be working for the first time in new spaces, collaborating with new team members, and may even be providing a type of care that is not their specialty (for example neonatologists caring for adult ICU patients). In medical decision-making, providers have been trained to weigh the level of risk against the potential for positive outcomes. But within COVID-19 units, the risk level of different spaces is not always clear or something clinicians consciously think about.

**DEVELOPING SPATIAL LITERACY**

Terms like spatial literacy (Tversky 2019) are underutilized in the research world but are especially prescient in a moment like this. As individuals, our spatial literacy around our bodies and our proximity to others has been unexpectedly heightened over recent weeks. The simple act of asking a clinician to annotate a heatmap by color-coding spaces to infection risk was revealing. The exercise generated interesting discussion and helped clinicians deeply consider space in a new way and rethink existing assumptions. The heatmap created for the pediatric COVID-19 unit also made visible an important design challenge: controlling movement across units. Particularly in tight, vertical healthcare settings, like the Mount Sinai Kravis Children's Hospital, healthcare teams needed to leave units in order to access key shared spaces like nurses lounges, or fetch resources like drugs or supplies. To limit nosocomial infection between COVID-19 and non-COVID-19 units, thresholds between floors (i.e., elevators and stairs) needed to be considered more closely, but proved difficult given the constraints in this building.

Figure 12: Annotated floorplan of pediatric COVID-19 unit by Mount Sinai clinicians.

Annotated heatmaps of the adult ICU units revealed significant differences in individual perceptions of risk zones. This is notable because staff behavior is typically most stringent in areas perceived as red zones, where the risk of contagion is perceived as high; but there were varying degrees of adherence to infection control rules in spaces perceived as orange or green, and differing levels of legibility of which spaces carried high risks of contagion.

Because of this ambiguity, "warm zones" may be dangerous for healthcare workers simply because they are not aware that they are warm zones. In warm zones, healthcare workers might overlap with staff coming from or going to other floors, or on-call clinicians might be changing personal protective equipment (PPE).

At a bare minimum, health facilities can leverage design cues to clearly mark orange areas as "warm zones" and convey expected behavior to healthcare workers within these areas. At the same time, they can implement active interventions to try and lower infection transmission in these areas—that is, turning orange spaces into green ones.

We asked three different clinicians to annotate heatmaps of ICU units on the same floor (Figure 13). Their drawings revealed three very different perspectives on risk.

- "Hot" areas known to be contaminated, where providers should use full infection control precautions and be on alert (red color in drawing).
- "Clean" areas that are fairly confidently non-infected areas, where staff can use relaxed precautions (green color in drawings).
- "Warm" or unclear areas that may actually pose risk even if not perceived as a risk zone (orange color in drawings).

Figure 13: Heatmaps of ICU units on the same floor by Mount Sinai clinicians.



Donning/doffing zone for units

Figure 14: Risk zones of an adult ICU floor during pandemic surge conditions, sketch shows the zones in the Med-Surg and ICU units on 6th floor of Guggenheim Pavilion, image copyright MASS Design Group.

Together in conversation with staff at Mount Sinai, the research team identified red, orange, and green zones under pandemic surge conditions. In the condition of a massive increase in infection control rooms, the entire unit effectively becomes an orange zone, and clearly marked donning and doffing areas need to exist at major entrances to the unit, in addition to patient room thresholds (Figure 14). All possible measures should be taken to ensure that stairs and elevators in vertical hospitals stay uncontaminated, as they connect many different zones; and tape, signage, or paint can be used to designate these different risk areas and thresholds more clearly.

**CONCLUSION**

While revised infection control practices have been quickly adopted, medical facilities are not configured to adapt easily to these rigid adherence protocols. Instead, they are forced to respond "on the fly" in situ, often with insufficient amounts of equipment, protective gear, or clear guidelines. Amidst heroic efforts to transform hospital-wide clinical processes

and infrastructure, adaptations can vary floor to floor and unit by unit, often revealing design "hacks" that are both creative and discordant. Nevertheless, hospitals are not designed for pandemic surge, and existing infection control protocols are inadequate for COVID-19.

**KEY TAKEAWAYS**

More research and strategies in partnership with healthcare staff and designers are needed to help develop spatial literacy among providers and to use the built environment to facilitate this knowledge. Thoughtful spatial interventions can help create situational awareness for COVID-19 units, set clear and consistent protocols for thresholds, align appropriate behavior within well demarcated risk zones, and identify opportunities to more safely and flexibly expand capacity during surge conditions.

1. While infection control protocols are being quickly adopted, hospitals are not designed to easily pivot to support the infrastructural changes needed at the scale of a pandemic surge. One major shift is the focus from individual patient care to unit care, creating caregiving processes for COVID-19 units that place a heavy reliance on interstitial spaces. This lesson means that enforcing proper infection control protocols in hallways and key thresholds, like entries into units, may be as important as in patient rooms.
2. Variability exists both in the adaptations from floor to floor and unit by unit, as well as in personal perception of risk zones within care units. These differences affect how health workers behave and interact, sometimes introducing new elements that may impact the risk of infection among staff and patients.
3. Spatial literacy (i.e., the ability to read and understand space) can be a powerful tool to orient healthcare workers within unfamiliar and rapidly-changing COVID-19 caregiving environments. Simple visual aids and design nudges can help mitigate infection transmission by clearly conveying risk zones, creating mental "anchors" for specific activities, and aligning behavior with policy. Design can help advance the sociology of infection control, along with our developing understanding of its biology.

These interventions are needed now, but they will not be short-lived. And they will be even more necessary in hospitals with fewer resources. Thoughtful and simple spatial adaptations can help create situational awareness for COVID-19 units, set clear and consistent protocols for thresholds, align appropriate behavior within well-demarcated risk zones, and identify opportunities to more safely and flexibly expand capacity during surge conditions.

As we continue to realign care infrastructure and processes to the new normal, clinicians, designers, and administrators must work together to make invisible risks more visible, and create spaces that instill order, collaboration, and morale.

**Michael Dolinger**, MD, MBA is currently the chief fellow of pediatrics at Kravis Children's Hospital at the Icahn School of Medicine at Mount Sinai where he is also a fellow in the department of Pediatric Gastroenterology and Hepatology. He has previously performed

successful research on physician and trainee wellness and burnout, and social network analysis among healthcare teams.

**Ashley Marsh**, RA, is a licensed Architect currently with MASS Design Group, where she's advancing COVID research in health, culture, and education settings for organizations across the US and generally leads strategic development along with consultative pre-development and readiness projects. She has deep experience conducting research in the built environment and designing user experience strategies for large, complex organizations, including previously leading research for San Francisco International Airport's Revenue Enhancement and Customer Hospitality (R.E.A.C.H.) initiative. She has been named '40 under 40' by the San Francisco Business Times, an Emerging Leader from the Design Futures Council, previously presented her research at forums like EDRA and SCUP, has held teaching positions at the Hasso Plattner Institute of Design at Stanford University and the Boston Architectural College, and was awarded the $10 million XQ Super School Prize for her research work on the future of the American public high school in Oakland, CA.

## NOTES

## REFERENCES CITED

Edmondson, Amy C., and Jean-François Harvey. *Extreme Teaming: Lessons in Complex, Cross-Sector Leadership*. Emerald Group Publishing, 2017.

Liu, Y., Ning, Z., Chen, Y. et al. "Aerodynamic analysis of SARS-CoV-2 in two Wuhan hospitals." *Nature* 582, (2020) 557–560. https://doi.org/10.1038/s41586-020-2271-3

Tversky, Barbara. *Mind in Motion: How Action Shapes Thought*. Basic Books, 2019.

Zimring CM, Matić Z, Wong Sala MF, et al. "Making the invisible visible: Why does design matter for safe doffing of personal protection equipment?" *Infection Control and Hospital Epidemiology* 39, 2018: 1375-1377. https://doi.org/10.1017/ice.2018.206

CATALYST

# From the Space Station to the Sofa

## Scales of Isolation at Work

JO AIKEN, *University College London/ETHNO-ISS (NASA)*
ANGELA RAMER, *HKS Architects*

*Since March 2020, many employees around the world have been forced to work from home due to the COVID-19 global pandemic. Astronauts aboard the International Space Station (ISS) have experience in working in isolation and confinement. This paper focuses on a comparison of astronauts on the ISS and Earth-bound architects and interior designers restricted to working from home (i.e. their sofas) due to the pandemic. Isolation at work emerges as a complex phenomenon characterized by the measured and perceived distances between physical, social, and temporal spaces. By examining the scale-making activities of NASA and HKS, analogs provide a possible means for studying and predicting the complex dimensions of isolation. The work ecosystem is a useful tool in conceptualizing and operationalizing the employee experience to design the future of work and workspaces.*

In 2020, the global workforce has become distanced in ways that no one could have imagined. Due to the COVID-19 pandemic, humans around the globe transitioned nearly overnight to working in isolation, confined to our homes, and conscious of the extreme (i.e. contagious) world outside. Leaders and employees in governments, businesses, universities, and organizations of every type are affected by the current pandemic. The sudden, worldwide shift to remote working has prompted many questions about the future of work (FOW) and the concept of work-from-home (WFH). Inspired by the messages from astronauts on the International Space Station (ISS), we consider the 2020 employee experience in two seemingly different industries, space exploration and architecture. Our focus is on a comparison of astronauts on the ISS and Earth-bound employees restricted to working from home (i.e. their sofas) due to the pandemic. The employee experience of working in isolation is contextualized through examples of our respective ethnographic work on workplace design at NASA on the design of space habitats and with HKS Inc., a global architecture firm. From the space station to the sofa, we explore the usefulness of the ISS as an analog[1] for the WFH experience and open up questions of scale, space, place, and time.

## ASTRONAUTS AND ANTHROPOLOGISTS: IN THIS TOGETHER

"From up here, it is easy to see that we are truly all in this together. #EarthStrong," astronaut Jessica Meir tweeted on March 16, 2020, from the ISS, shown in Figure 1. As COVID-19 spread across the globe, the first messages from astronauts focused on human solidarity. Astronauts are known to share Tweets and words of encouragement during natural disasters, terrorist attacks, and other catastrophes that affect nations around the world. COVID-19 was no different. However, with the sudden shift to remote working occurring worldwide, astronauts began adding words of wisdom to their messages from

outer space. Medical experts and government officials began directing the world on how to live during a global pandemic, and astronauts quickly began sharing their expert tips for working in isolation. On March 21st, days after Meir's first Tweet, retired astronaut Scott Kelly (2020) published an opinion piece in *The New York Times* with the title "I Spent a Year in Space, and I Have Tips on Isolation to Share," see Figure 2. NASA (2020) followed suit the following day (Figure 3) and published tips from astronaut Anne McClain on its website. Two days later, the *National Geographic* (2020) published an interview with astronaut Chris Cassidy "Stuck in a cramped space? This astronaut has some advice," shown in Figure 4.



Figure 1. First astronaut Tweet regarding COVID-19

Figure 2. Retired astronaut Scott Kelly's (2020) Tips for Working in Isolation



Figure 3. NASA (2020) posts astronaut's tips for working in confined spaces

From the Space Station to the Sofa – Aiken & Ramer

Figure 4. Astronaut Chris Cassidy shares tips for working in confinement (National Geographic 2020)

What started as texts between friends, we came together intrigued by the number of interviews, articles, podcasts, and social media posts of, by, or from astronauts and cosmonauts giving WFH tips. As intrigue grew towards intellectual curiosity, we began reflecting on anthropological notions of isolation, work, and scale. As two applied anthropologists from two different industries, we also drew on our respective work experiences and ethnographic research. From a space anthropologist's perspective (Aiken), how does isolation at home compare to isolation in extreme environments? As a design anthropologist working in architecture (Ramer), how alike is the work-from-home employee experience to that of an astronaut? For both of us, as applied, design ethnographers, these questions converge on the subject of scale. How far down, or up, can you scale human experiences of work in isolation? In other words, how comparable are the astronauts' work experiences in isolation to employees working at home during a global pandemic? How useful would such a comparison be? Can you measure, or scale, isolation and work

experiences therein? How do these scales develop in the first place, and how useful are they when it comes to designing the future of work, in space and on Earth? The discussion that follows unpacks these questions and more.

## ISOLATION AT WORK

At first glance, working in outer space seems very different than working from home on Earth; and at the same time, the astronauts' tips for working in isolation during COVID-19 are worthy of international media attention. *Isolation* is the common element that connects these two environments of work in 2020 (i.e. the sofa and the ISS). Simply defined, *isolation* is the condition of being isolated or "set apart from others" (isolate, n.d.). Anthropologists, psychologists, and other related theorists offer insights on what it means to be "set apart" through discussions of space, place, and time (see Low and Lawrence-Zúñiga 2003). Aiming toward practical design application, isolation can be grouped into two categories, or dimensions: *physical space* and *social space*. One can be physically set apart from others as well as socially set apart. Applied, ethnographic workplace studies often address aspects of physical and social spaces (see Cefkin 2010 and Gunn, Otto, and Smith 2013). However, time is a dimension of isolation that is less discussed in anthropology and design research regarding work.

### Physical Space in the Workplace

A practical, even mundane, view of *physical space* focuses on the visible, measurable distances between objects and people in enclosed environments. Physical space can be translated into the volume of a structure and/or the surface of an area. The physical environment and its boundaries are ultimately experienced and evaluated through the body's senses. Bourdieu's (1977) *habitus* suggests the body inhabits an environment that imposes structural constraints, forming dispositions or schemes of perception or thought. Workplace theorists have demonstrated that the perceptions of the physical environment, or office space, directly affects job attitude and performance (Kupritz 2000). Beginning with the first "modern" office space design in the early 19th century, American engineer Frank Taylor sought to maximize efficiency and productivity by designing workplaces based on the design of factories (Kupritz 1998, 2011). Taylorism evolved toward human relations and eventually toward more human-centered design practices in which workplace designs became more individualized and flexible. Over the years, trends in office design have fluctuated from private offices, to open floor plans, to cubicles, benching, assigned seats, hoteling, and hot desking. In outer space, workplace designs are much more limited. However, the design of the ISS also designates specific spaces for work activities that are physically separate from living quarters. As EPIC contributors Imai and Ban (2016) state and most workplace theorists agree, "the physical distance between workers has dramatic impact on productivity and collaboration." The physical environment, then, is a crucial component of understanding the total employee work experience.

## Social Space in the Workplace

Like astronauts, many 2020 WFH employees do not work in complete autonomy. Employees work as part of an organization, department, or team; they have co-workers and a boss. Following our simple view of physical space, *social space* can be described as the perceived distance between people in a given environment. However, unlike physical spaces such as private offices and open floor plans, social spaces are often invisible and immeasurable. As mentioned above, designers consider the physical spaces between members of an organization or team in designing office spaces that promote collaboration and productivity. In other words, physical spaces influence or create boundaries for functional social spaces in the workplace. We create social space, or separate ourselves, from others on the subway by wearing headphones and closing our eyes. The WFH employee appears isolated when his toddler is taking a nap, and no one else visibly or audibly makes a surprise appearance in the virtual meeting. Add in the ancillary layers of the digital realm *and* the added layer of time and we see an even more dynamic view of social space. The act of working outside of standard work hours is an example of creating social isolation, often at the excuse of needing privacy for more focused work. Increasingly, with the invention of the internet, social space has evolved to include virtual spaces that are not geographically based or digitally fixed in time (Miller and Slater 2000). Boellstorff (2015) has taught us to avoid framing the digital or virtual world as "unreal" in contrast to the physical world. Social space, then, exists on measurable scales of the physical as well as in the digital realm of connectedness. When considering the current COVID context, the varying states of social connection in terms of isolation comes to the forefront.

## Temporal Space (or Space in Time) in the Workplace

Time is a complex phenomenon. Here, we refer to *time* as a point of existence measured in units (e.g. hours, minutes, seconds) relative to a given standard. The time in San Francisco is 9:30 AM in the Pacific Time Zone while the time in London is 6:30 PM according to the Greenwich Mean Time (GMT) zone. For this discussion, we consider *temporal space* as the difference in time (as measured, perceived, or experienced) between people and events. At 9:30 AM Pacific, the architect in San Francisco has just begun her workday while, concurrently, the design researcher in London is preoccupied with preparing dinner. In this example, the architect and the design researcher can be engaged in the same virtual meeting, but their experience is different due to the measurable, temporal space between them. Time, though addressed with less frequency and less explicitly in applied workplace research, is a key contextual factor in working in isolation. Astronauts understand that their job requires working in isolation at some point in their careers – once assigned to a mission, they know the start date (i.e. time) their isolation begins. It is unlikely that many architects or designers knew a global pandemic would change their work environment before it happened in 2020. Isolation in spaceflight also has a confirmed end date. Space missions, like polar expeditions, are planned – astronauts know when they are going home. Antarctic scientists say that knowing the last day of the mission was sometimes what got them through the loneliness. In a 2013 habitat study, one NASA scientist shared with Aiken, "If for some reason, your 'Going Home Day' changes... some people lose their minds over a 3-day delay." The scientist went on to emphasize that a change in the extraction date is just as meaningful and

often more impactful on mental health than an extended duration. For many Earth-bound workers, at the time of this writing, it is still unknown when offices will resume normal operations. With no end date and no set duration, the total impact of working in isolation at home during COVID-19 remains unknown.

## EXPLORING THE USE OF SCALES: ANALOGS OF ISOLATION

Space scientists and engineers as well as architects and interior designers create and use scales to address the physical, social, and temporal dimensions of workspaces. As seen through two work examples, HKS and NASA experts create and use scales to make tangible and design for the unknown. Tsing (2011:57) uses "scale-making" to refer to projects that create or transform the perception of a scale (i.e. the global) to see how it might work on another scale (i.e. the local). NASA engages in scale-making through the use of local (i.e. located on Earth) space analogs to imagine life on a larger, galactic scale (i.e. not located on Earth). Similarly, HKS views its local architects in a scale-making effort to understand the global WFH workforce – the 2020 sofa is, in a way, an analog for understanding the future workplace experience. NASA and HKS seek to understand and make sense of the dynamic nature of physical, social, and temporal spaces in an isolated workplace. An analysis of these analogs is crucial in assessing the possibility of scaling up the experience of the ISS astronaut to the global, isolated 2020 workforce.

### Designing Habitats for Space Exploration

As an applied researcher internal to NASA, Jo Aiken conducted astronaut workplace studies from 2013 to 2018. The first project involved the design of habitats to be used for future missions to Mars. Habitat design is a unique, complex challenge for NASA's engineers. As humans explore space beyond low-Earth orbit (LEO) new habitability concerns emerge that influence the design of a habitat. Sending humans to work off-planet is costly, so the design requirements also take into account the cost of launching a heavy space habitat. When making critical decisions, space architects and engineers design space systems based on legitimate requirements and not what is simply "nice to have." To assess habitat design considerations that are more than "nice to have," Aiken gathered ethnographic evidence about perceptions and behaviors related to living in a Mars habitat. The study provided meaningful insights into previous assumptions made by engineers such as the importance of designing separate spaces for working, eating, sleeping, and exercise. The study also resulted in 25 key findings relating to privacy. Several of these findings provided strong evidence for designing private crew quarters larger than what is currently provided on the ISS – privacy is more than a "nice to have" even when living and working in an environment of isolation. NASA engineers continue to develop the design requirements for space habitats based, in part, on these findings.

Since humans have yet to live on a celestial body other than Earth, it is difficult for engineers to contextualize interactions between astronauts and the technology that is required for living and working on another planet. Aiken, like other NASA researchers, used space analogs to study what it would be like to live in a Mars habitat. Space analogs include simulated missions to Mars as well as science research stations in Antarctica and submarine crews. They use space analogs to scale-back the level of difficulty in studying a future, off-world workplace. Space analogs are located on Earth, yet they mimic to some degree the experience of working in outer space. NASA engineers and scientists refer to the environment of space as Isolated, Confined, and Extreme, or simply "ICE." NASA researchers use Earth-based analogs to study what it is like to live and work in an ICE environment. They study scientists wintering over in Antarctica to understand isolation and sensory deprivation. They learn about living in small, confined spaces from submarine crews. While nothing can exactly simulate living and working in space, space analogs and astronaut-like populations are characterized by their degree of similarity to the space ICE environment. For example, a simulated Mars mission on the Big Island of Hawai'i is a high-fidelity analog due to the isolated and confined conditions of the participants living in a small habitat. A group of scientists wintering over in Antarctica is a higher-fidelity analog because of the increased remoteness and extreme environment of the region. The ICE scale, on Earth as well as in space, is fluid and changes due to the weather, the sound of a tourist helicopter flying over Mauna Loa, or constant video monitoring by a simulated mission control. NASA uses the concept of ICE to study, predict, and plan for the dynamic nature of social spaces in outer space - ICE influences the construction of space structures, or physical space, as well as the social aspects of space exploration (NASA 2014).

NASA uses another scale to plan for habitat designs at a more micro level than the ICE scale affords. In 2013, NASA asked its scientists and researchers to determine the minimum Net Habitable Volume (NHV), or the minimum number of cubic meters/feet, necessary for supporting crew life on long-duration exploration missions. The NASA behavioral health scientists and human factors engineers recognize the complex, dynamic nature of habitable volume. NHV is what is left available to the space crew after accounting for elements that decrease the usability or functional volume of the spacecraft. For example, the stack of books left in the backseat of my Jeep decreases the NHV available to my passengers. On Earth, gravity reduces the available functional NHV in a given workplace. An interior designer cannot, without assistance and potential harm, work on his ceiling. Astronauts on the ISS can utilize all four walls, or boundaries, of their physical space. In this way, physical space is scaled down on Earth – our available, functional physical space is limited compared to that in outer space. However, NHV or physical space is also easier to scale up in an Earth-bound workplace. A biomedical scientist can go for a walk outside the lab to create physical distance for privacy. The trash is cleared regularly. In space, emptying the trash and creating physical space is more difficult.

## Designing Workspaces for the Future of Work-From-Home (WFH)

Working as an in-house researcher for the global architecture firm, HKS, Angela Ramer works alongside architects and interior designers in the design of commercial workspaces on

Earth. Projects seek to scale the human/machine elements to appropriately create the best, most functional built environment and employee experience. This scaling of space for function is known as *programming* in which critical needs are identified and outlined in order of magnitude (e.g. high-level components) or are more detailed (e.g. an itemized list of dimensions, spaces, etc.). Additional affordances are applied for things like assigned seating, meeting/collaboration seats, shared amenity capacities (e.g., cafeterias, auditoriums), as well as code-compliant affordances like parking spaces, distances to entrances/exits, and elevators. Scales, or measurements, can also relate to headcount and occupancy (e.g., building and room capacities) and seating assignments (e.g., individual desks/offices or shared workstations/offices). The most common spatial measurements include Gross Square Footage (GSF), Rentable Square Footage (RSF), and Useable Square Footage (USF). These are interdependent scales used to describe and determine the appropriate allocation of space the results of which are intertwined with facility operations, business goals, human experience, and organizational culture. Similar to NASA's use of NHV, these scales are particularly used concerning function. COVID-19 presents the opportunity for a transformational change in the way architects and interior designers think through and engage with the scales in the post-COVID future workplace.

*HKS & The Sofa Analog*

As NASA uses space analogs on Earth to test and train for missions in outer space, HKS is in a way engaging the "sofa" (i.e. the home) as an analog to study the office of the future. This approach aims to generate insights to inform organizational operations, real estate, and employee work experience decisions in light of the COVID-19 era workplace. Since March 13th, Ramer has been supporting various HKS research initiatives to study the employee experience across the globe. HKS sends surveys at key intervals to track employee experiences regarding mental health, social connections, environmental conditions, work processes, and more (see HKS 2020b). Ramer and her colleagues tracked responses to a core set of questions over time (e.g., desire to continue working from home, activities best done in an office environment, work-life balance, fatigue, etc.). This data was augmented with timely, topical survey questions (e.g., satisfaction with return-to-work protocols, satisfaction with home-work environment). The data from the COVID-era surveys is triangulated with employee surveys collected pre-COVID. Through this continuing process, Ramer and her colleagues seek to identify the social and functional affordances of home-work environments and lifestyle factors previously considered irrelevant in relation work environments. The purpose is to uncover the role that home now plays in the employee experience and to what extent employers need to adapt their policies and spaces to support this fundamental shift in how work gets done.

Findings from the recent HKS studies highlight the significance of work activities within varying scales (i.e. measurements) of connectedness (i.e. isolation concerning social spaces) and the effects of the workplace on our overall health. A factor analysis of the survey data reveals that only two employee demographics are significant predictors of the overall WFH experience – employee age and housing type (e.g. single family home, apartment, etc.) (HKS 2020a). These individual attributes previously regarded as outside the realm of employer consideration are now at the forefront in considerations of organizational policy, culture, and real estate. These attributes have also been found to directly impact the ways and to what

extent isolation is experienced in the WFH context. The HKS studies have found that a sense of connectedness is lower for those who live alone. Likewise, living alone is an indicator of a higher desire to return to the office (HKS 2020a).

In industry conversations, there are many terms related to and sometimes used interchangeably with social *isolation: separation, segregation, seclusion,* and *insulation,* and more recently *social distancing,* and *quarantine.* Architect and design researcher, Erin Peavy (2020) explores important differences:

> Although *loneliness* and *social isolation* are often used in the same breath, the two are distinctly different. Loneliness is essentially the perception of social isolation, whereas social isolation is the absence of regular human interaction in one's life. These phenomena are tied to belonging, trust, social cohesion (the strength of the bonds among members of a community) and social capital (the tangible and intangible benefits a person reaps from his or her social network) as components of our social health, defined as a critical aspect of overall health.

The effects of loneliness and social isolation on occupational health are exacerbated by the current COVID climate where many are still relegated to working from home while many are also living alone. More than 60% of home workspaces are not dedicated or designed as home offices. Employees work from sofas, kitchen tables, bedrooms often alongside their children attending school from home (HKS 2020a). HKS seeks to understand the complex variations of home-work environments, where employee needs are being met, and what is lacking so that employer-provided workspaces can be redesigned to create the best remote and co-located employee experience possible (HKS 2020a). To fully scale up the sofa analog, HKS intends to engage with clients in other industries, share initial insights, and expand data collection to inform the design of future workplaces for other office-based professionals.

## Comparing Contexts: The Space Station and the Sofa

The functional differences between scales of physical space in the ISS workplace and the terrestrial workspace are largely driven by gravity and the harsh environment of outer space. While we are physically isolated, or separated, from others by the walls and doors of our homes and offices, a critical point of differentiation is the context in which these scales of isolation are being experienced. Living and working in microgravity is a complex practice that requires years of planning. For astronauts, working in isolation is their desirable end-goal achieved through years of training. Astronauts are not thrown into isolation, or microgravity, on their own. They are assigned to a mission crew, training as a team to minimize the risks of working in an isolated environment. In a 2015 NASA technical report, psychologists emphasized the continued use of training aimed at developing resilience to isolation in crewmembers (Vanhove et al. 2015). NASA provides isolation training through various means, one of which is by sending astronauts to its underwater analog, "NEEMO." They start slowly – NEEMO expeditions last only up to three weeks. NASA also trains its non-astronaut employees. Less than a month before the nationwide stay-at-home orders due to COVID-19, NASA held a mock "stay-at-home day" for its employees. Unlike NASA employees, many 2020 stay-at-home workers did not have the opportunity to set up home offices before orders were in place. Few, if any, spent years of training for working in isolation. For most Earth-bound workers, WFH orders have been unexpected and in many

cases undesired state with the little-to-no period of preparation or training. Table 1 shows a brief comparison of working in isolation during COVID-19, contextualizing the astronauts on the ISS and the WFH experience of architects and interior designers.

Table 1. Brief Comparison of Working in Isolation During COVID-19

| | NASA Astronauts on ISS | Architects/Designers at Home |
|---|---|---|
| **PHYSICAL:** **Location of work** | Low Earth Orbit (LEO) Space station habitat | Earth-bound Personal homes |
| **SOCIAL:** **Work + life** | Individual contributors Co-located crewmembers Remote Mission Control teams<br><br>Remote family, friends | Individual contributors Remote small teams Remote large teams<br><br>~Co-located family Remote family, friends |
| **TIME:** **Zones, Scheduling of work + events** | Operates on GMT/UTC, Coordinates activities on any given Earth time zone<br><br>Experiences 15-16 sunsets every day | Operates on time zone relative to individual location, Coordinates activities with selected co-workers in various time zones<br><br>Experiences 1 sunset every day |
| **Type of Work** | Scientific experiments/Research Station maintenance Public outreach | Design work Analytical Administrative tasks |
| **Skills/Training** | Selected for STEM education and physical fitness 2 years (avg.) of Astronaut Candidate (ASCAN) training +6mo. Mission-specific training | Hired based on architecture education and experience<br><br>Continued education for licensing |
| **Tools** | Highly specialized equipment designed for microgravity General hardware/software | General hardware/software Assigned, general equipment |
| **Duration (vs. Time, as denoted above)** | Intensive, short term missions (currently 6months – 1 year) | WFH efforts, duration currently unknown |

## ISOLATION AND SCALE IN THE FUTURE OF WORK

Although the ISS represents an extreme and unique case of isolation in the workplace, scales of isolation in the outer space workplace are useful in thinking through the socio-spatial challenges of working on Earth. Anthropologists and other social and behavioral

theorists can contribute to a greater understanding of isolation at work. Designers with a deeper understanding of these notions can design better workspaces, in space and on Earth. As early as the 1980s, design researchers understood the importance of situating product use in its sociocultural context (Wasson 2000). Buchli (2013), through an anthropological view of architecture, explains architecture as something other than a static space – architectural spaces sustain, shape and re-shape, social relations. So, what sociocultural knowledge leads to a greater understanding of isolation in the context of work? Most importantly, how do we connect a greater understanding of social space, or social isolation, to a practical approach for designing physical workspaces?

Physical isolation, although fluid and dynamic, is considered by NASA and HKS as something visible and measured. Astronauts are physically separated, isolated a measurable distance away from Earth. During COVID-19, architects are physically separated from their peers, working in isolation from their homes. Social isolation, as an absence of human interaction, is much more complicated. Are we ever truly socially isolated? Strathern's (2005) *merographic connections*, a way of knowledge-making that considers things as always part of something else, is particularly useful in exploring social isolation in the workplace. Stay-at-home COVID-19 workers are separated from traditional, face-to-face interactions with their co-workers. Astronauts onboard the ISS interact with their crewmembers, but they are isolated from interactions with the NASA workforce at large. These interactions, or perceptions of, are mediated through the use of technology. Mission control and Zoom meetings maintain a level of connectedness between the physically isolated workforce. According to Strathern's idea of merographic connections, this dynamic nature of separate-yet-connected occurs simultaneously. Individuals appear separated, socially isolated, from one point of view. At the same time, they are also connected from another point of view. The separate-yet-connected worker is simultaneously part and whole. Therefore, **isolation is a situated concept**.

So how do we situate this complex concept of connectedness (vs. isolation) in a physical, workplace architecture? Theoretical physicist and feminist theorist Karen Barad offers additional insight into isolation beyond a simple absence of human interaction. Barad (2007) coined the term *intra-action*, as opposed to interaction, to describe the agency of people, nature, and 'things.' *Interaction* presumes that when two entities come together they maintain a level of independence. *Intra-action*, on the other hand, suggests that entities act in co-constitutive ways – their agency is not a preexisting given. In simpler terms, actions are a result of relationships. Following this school of thought, an individual working in complete social isolation is impossible because "individuals" or entities do not have agency outside a particular intra-action. Furthermore, entities that come together to intra-act do not have to be human. The lone artist intra-acts with paint, brushes, and a canvas to produce a work of art. COVID-19 is an intra-action between human and non-human actors; the global pandemic has agency because of these intra-actions. Astronauts work onboard the ISS because of their intra-actions with their Earth-bound co-workers. An employee cannot work in complete isolation, therefore, because **actions are situated in relationships**.

So far, we have explored well-respected, yet abstract theories to breakdown the concepts of isolation and work. From Strathern (2005), we learn that isolation is a situated concept. Barad goes further to explain that work, or actions, are situated in relationships and individuals cannot act in isolation. As we look to connect social isolation with the workplace environment, Edward Hall's (1966) theory of *proxemics* provides a tangible, body-centric look

at perceptions of space and workplace needs. The four types of distances people keep (intimate, personal, social, and public) are learned through observation. Developed over fifty years ago, Hall's study of how humans use space within the context of culture is still useful in the design of built environments. Microsoft's Caitlin E McDonald (2020), a digital anthropologist, noted in a recent article that "the communicative aspects of proxemics are very important as we consider the ongoing disruptions to working and living as a result of the pandemic." Significant to the WFH worker, digital proxemics considers uses of physical and virtual spaces in connection with the uses of technology. McDonald (2020) suggests that organizations should consider replicating Hall's proxemic zones when communicating virtually. Communication, then, is a result of the WFH intra-actions of people and technology. Virtual meetings as well as the physical office produce workplace relationships. In other words, **relationships are facilitated through an environment**.

As we move closer to connecting social isolation, or the lack thereof, to the workplace environment, it is important to take a step back and consider what we observed as practitioners at NASA and HKS. In the habitat study example as well as the WFH architect survey, designers in both fields rely on scales to make sense of the work environment. The scales they use are dynamic, suggesting the ever-present possibility of change. McCabe and Briody's (2017) recent engagement of *assemblage theory* explicitly addresses the complex nature of change in organizations. Assemblage theory, first developed in the 1980s by Deleuze and Guattari (1987), provides a framework for analyzing social complexity as fluid and temporary. Assemblages are comprised of people, things, practices, discourses, organizations and institutions, and the complex nature of these components means that change is inevitable. **Relationships are situated and facilitated in an environment of constant change**.

In summary, isolation and work are actions that are situated and facilitated through relationships that exist in an environment of constant change. Viewing the workplace as an ecosystem, an emerging concept being developed from the HKS WFH studies, provides a means for grounding this complex notion in practical applications for designing the future of work.

## THE WORK ECOSYSTEM FRAMEWORK

The *work ecosystem* framework (see Figure 1) brings together Gibson's concept of *affordances* (1966) and McCabe and Briody's *assemblage theory* (2017) to capture the more holistic picture of what WFH looks like in a COVID and post-COVID world. It reflects more fully on the work experience to include place and process with underlying layers of process, time, and technology. This framework prompts a paradigm shift away from independent employer and employee context into an integrated and interdependent relationship. This interdependent nature mirrors the more intense alignment between astronauts, their environments, and their mission-critical survival. Thus, it is less about where and when we work as disparate attributes but more about how we work that ensures success. For example, Earth-bound workers have shifted from work-from-home being an ad-hoc, office alternative environment- often unavailable due to organizational policy, workflow or position (e.g., only mid or senior level staff had approval), or an employee's circumstances (e.g., lack of effective workspace, poor home internet bandwidth, etc.). The *work ecosystem* framework also taps into Hall's notion of *proxemics* (1966), with the blurring of

both physical distance and social space into scales of perception and experience; however, the *work ecosystem* acknowledges the interdependent but not necessarily nested attributes of space.
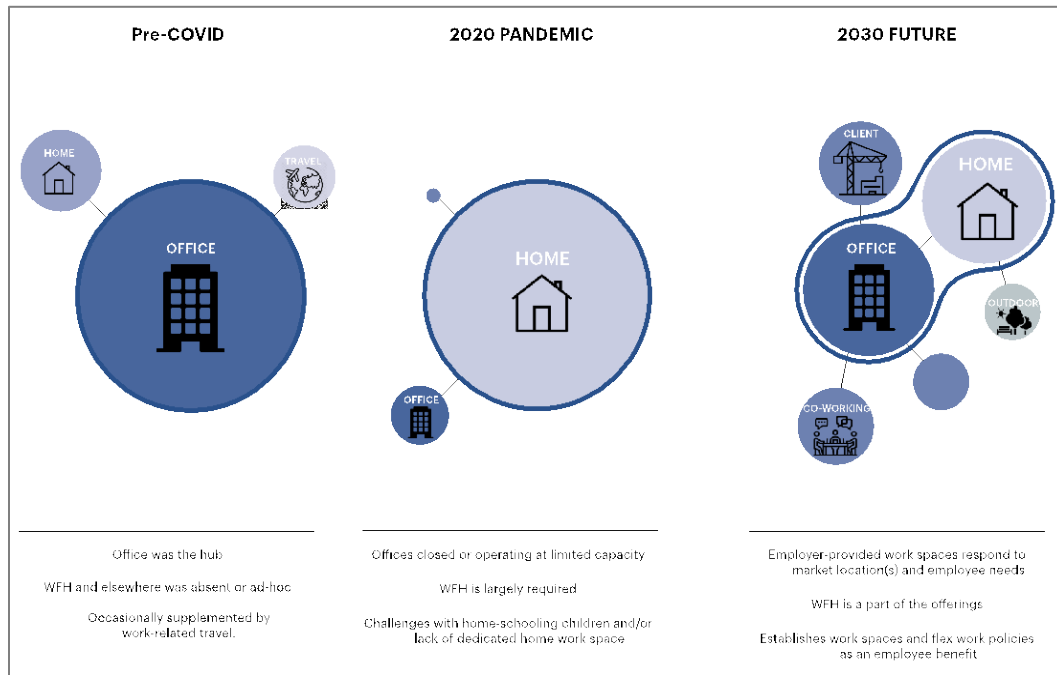


Figure 5: Work Ecosystem Framework (From HKS, pending HKS publication, used with permission)

In the current COVID climate, where home remains the primary work location, the reliance on a binary work system (e.g., home and office) serves as a distorted view of work. One life where they work and works where they live- a unique, integrated existence known well to those on the ISS. In considering the *work ecosystem* framework and the *assemblage* approach in a post-COVID time, organizations can value the role of multiple environments at varying scales for both the employee (spaces available to them) and employer (spaces offered). This has a substantial impact on organizational real estate (from consolidation/downsizing and campus planning to rethinking the need for single-tenant space and considering workspace alternatives more seriously, e.g., co-working memberships, subsidizing home office environments). Many of the astronaut-recommended adjustments appealed (and in many cases were accurate) for the initial adjustments to working from home (e.g., establish a routine, build in a mental commute, find/get a hobby). However, their real value was in offering a pragmatic crash course based on personal experiences to help with the short-term shock of WFH, especially WFH during a pandemic. They were quick fixes in an era of evolution. Six months later, while they fall short of formalizing the larger transformational changes that office work is looking for, they have provided perspective as organizations rethink and rebuild a framework for work. One with an expanded appreciation for affordances, with appropriate training and resources- giving office workers less of a do-

it-yourself survival guide and more of a foundational set of work in isolation attributes that can then inform their decision to continue (or not) in a WFH setting.

## CONCLUSION

COVID-19 has served as an unavoidable catalyst for the evolution of work- one we could have never expected. This goes beyond another iterative of the decades' long debate between topics of private or open offices, cubicles, hot desking or hoteling, standard business hours and flex work arrangements. WFH in COVID times has ushered in a fundamental paradigm shift in affordances (and acceptance) for where, when, and how work gets done. Employees and employers are now connected beyond paychecks for services rendered or physically populating real estate. They are interconnected ecosystems with elements operating at various scales. So, while astronauts were effective coaches in the initial transition to WFH, we've found that the nuances of isolation were best understood through the lens of *intentionality*. While romanticized and potentially limited in the civilian view of the role, astronauts know and accept the risks of their exploration-based employment- with its controlled projects, hyper-specific testing and retesting, simulations, and psychological support. The average office worker turned remote worker grapples with an entirely new work context with equally unexpected co-workers (e.g., spouses, children), little-to-no preparation, and every day feeling like it is all part of one big experiment with no end in sight.

The 'future of work' has been and will continue to be an ever-evolving state of being. Our reflective analysis serves as a snapshot of precedence and current context, seeking to inspire further dialogue during this transformational moment. Nuances between industries, geographies, and policies greatly impact organizations' abilities to provide for such complex considerations, however, it would be naïve to suggest that one could plan for every possible permutation. Instead, organizations need to consider their *work ecosystem* relative to that of their staff to make informed decisions for policies, processes, and place-based experiences. These reflections and recommendations leave us with several questions for the EPIC community to carry forward in the discussion and further exploration as the COVID-19 context evolves:

- How will post-COVID work experience impact commercial real estate and the continuation of offices as workplaces?
- In what ways will organizations (re)consider home office support (e.g., stipends for Internet, office supply subscriptions) as an employee amenity or even a necessary extension of the office environment for work?
- How best can organizations maintain, adapt, and embody organizational culture through digital presence?
- In what ways will WFH policies impact the recruitment and retention of talent? How might this relate to a rethinking of 'talent pools'? What could it mean for staffing?
- To what extent will the rethinking of work and workplace impact business and operating hours?

- How does an organization address technological competency in the current workforce as well as set expectations for future employees? What are the implications connected to higher education curricula?

**Jo Aiken** is an applied anthropologist with research interests and professional focused on the intersection of organizational culture, design futures, and innovation. Before joining UCL's ETHNO-ISS project, Jo worked in healthcare, aviation, and aerospace. She has over 20-years of experience at NASA working in various roles from Mission Control to human factors engineering and executive leadership consulting. (jonieaiken@gmail.com)

**Angela Ramer** is a design anthropologist at HKS Architects specializing in strategy, with a focus on the human experience within designed environments. Her latest work focuses on strategizing new healthcare service offerings, future-casting the role and design of corporate work environments, and reimagining the community spaces for engagement and connection. (aramer@hksinc.com)

## NOTES

1. This paper uses the spelling of "analog", "analogs" common within NASA.

## REFERENCES CITED

"isolate"
n.d.        Merriam-Webster Online Dictionary. Accessed July 10, 2020. https://www.merriam-webster.com/dictionary/isolate .

Barad, Karen
2007       Meeting the Universe Halfway: Quantum Physics and the Entanglement of Matter and Meaning. Duke University Press.

Boellstorff, Tom
2015       Coming of age in Second Life: An anthropologist explores the virtually human. Princeton University Press.

Bourdieu, Pierre
1977       Outline of a Theory of Practice. Volume 16: Cambridge University Press.

Buchli, Victor
2013     An Anthropology of Architecture. Routledge.

Cefkin, Melissa, ed.
2010     Ethnography and the Corporate Encounter: Reflections on Research in and of Corporations. New York: Berghahn Books.

Deleuze, Gilles and Felix Guattari
1987     A thousand plateaus: Capitalism and schizophrenia. Translated by Brian Massumi. Minneapolis: University of Minnesota Press.

Gibson, J. J.
1966     The senses considered as perceptual systems. Boston: Houghton Mifflin.

Gunn, Wendy, Ton Otto, and Rachel Charlotte Smith, eds.
2013     Design Anthropology: Theory and practice. A&C Black.

Low, Setha M., and Denise Lawrence-Zúñiga, eds.
2003     The Anthropology of Space and Place: Locating culture. Blackwell.

Hall, Edward
1966     The Hidden Dimension. Garden City, N.Y.: Doubleday.

HKS
2020a     "From Temporary to Transformative: Work from Home Insights."https://www.hksinc.com/how-we-think/research/from-temporary-to-transformative-work-from-home-insights, accessed 22 August 2020.

HKS
2020b     Work Beyond the 4th Dimension. Publication pending.

Imai, Ryoko, and Masahide Ban
2016     "Disrupting Workspace: Designing an Office that Inspires Collaboration and Innovation." In Ethnographic Praxis in Industry Conference Proceedings 2016(1): 444-464.

Kelly, Scott
2020     "I Spent a Year in Space, and I Have Tips on Isolation to Share." The New York Times. https://www.nytimes.com/2020/03/21/opinion/scott-kelly-coronavirus-isolation.html, accessed 15 July 2020.

Kupritz, Virginia W.
1998     Privacy in the work place: the impact of building design. Journal of Environmental Psychology 18(4): 341-356.

Kupritz, Virginia W.
2000     Privacy management at work: A conceptual model. Journal of Architectural and Planning Research: 47-63.

Kupritz, Virginia W.
2011     Individual and Group Privacy Needs Across Job Types: Phase 1 Study. Journal of Architectural and Planning Research: 292-313.

McCabe, Maryann and Elizabeth Briody, eds.
2017      Cultural Change from a Business Anthropology Perspective. Lexington Books.

McDonald, Caitlin E.
2020      "Pandemic-Informed Proxemics: Working Environment Shifts Resulting from COVID-19."
https://www.microsoft.com/en-us/research/uploads/prod/2020/07/NFW-42-McDonald.pdf,
accessed 20 July 2020.

Miller, Daniel and Don Slater
2000.     Internet. Berg Publishers.

NASA
2014      Human Integration Design Handbook (HIDH). Revision 1. NASA/SP-2010-3407/REV1.
https://www.nasa.gov/sites/default/files/atoms/files/human_integration_design_handbook_revisio
n_1.pdf, accessed 15 July 2020.

NASA
2020      "An Astronaut's Tips For Living in Space – Or Anywhere."
https://www.nasa.gov/feature/an-astronaut-s-tips-for-living-in-space-or-anywhere, accessed 15 July
2020.

National Geographic
2020      "Stuck in a cramped space? This astronaut has some advice." Interview with NASA
astronaut Chris Cassidy on March 24, 2020. https://www.nationalgeographic.co.uk/space/2020/03,
accessed 15 July 2020.

Peavy, Erin
2020      Connecting IRL: How the Built Environment Can Foster Social Health. March 23, 2020.
hksinc.com/how-we-think/research/connecting-irl-how-the-built-environment-can-foster-social-
health, accessed 20 July 2020.

Strathern, Marilyn
2005      Partial connections. Rowman Altamira.

Tsing, Anna L.
2011      Friction: An ethnography of global connection. Princeton University Press.

Vanhove, A. J., Herian, M. N., Harms, P. D., & Luthans, F.
2015      Resilience and growth in long-duration isolated, confined and extreme (ICE) missions.
NASA/TM-2015–218566. https://ston .jsc.nasa.gov/collections/TRS/_techrep/TM-2015-
218566.pdf, accessed 30 July 2020.

Wasson, Christina
2000      Ethnography in the Field of Design. Human Organization 59(4):377-388.

**THEMATIC SESSION**

# Overcoming

This session identifies three scale-related challenges for applying ethnography in business and distills practical lessons for overcoming these challenges. First is the challenge of initializing and sustaining ethnographic practice in large scale organizations. Second are the organizational and cultural barriers to scaling ethnographic research within firms. Third is the challenge of resolving the scale differences between qualitative and quantitative research when trying to integrate the two methodologies.

Session Curators: Evan Hanover, Thomas Lee, Elena O'Curry, Tabitha Steager

# Scaling is Like Making Sourdough

## Finding Sourdough Starters to Help Your Research Scale

KARYN GEORGILIS, *Harvard Business School*

*Customer ethnography and user research continues to move higher up the priority list of Fortune 500 corporations. As a design researcher at a global consultancy, my clients often consist of new or aspiring consumer research groups eager to scale quickly. Excited at first, these groups or individuals are ready to dive in but get discouraged by the size and price tag of "big leap user research projects" then end up never pursuing ethnography at all. Watching this pattern unfold client after client, it started to remind of making sourdough. Because novice bakers start out trying to make sourdough from scratch, expecting heaps of picturesque loaf of bread right off the bat. But that's not how sourdough is made. The first step is finding a "starter". Sourdough starters are small pieces of fermented dough that one can really only get from an experienced baker. You need to integrate it into your ingredients and to make the sourdough rise, scale, and bubble. The same is true of starting and scaling internal user research groups and initiatives. So what's the equivalent to a sourdough starter? What are some "scaling starters"?*



Sourdough ingredients spread haphazardly across a kitchen counter. Unnamed by Anshu A. (CC BY 2.0)

**Karyn Georgilis** is a Design Researcher who's helped organizations marry consumer and commercial insights at the Mayo Clinic's Center For Innovation, REI, the MIT Media Lab, and various consultancies.

CASE STUDY

# DIYing along with DIYers:

## Juggling with Scales During Home-Improvement Research

GUILLAUME MONTAGU, _unknowns

*As a team of researchers was asked by a French home-improvement retailer to redefine their strategy, they designed and carried out an ethnographic and quantitative research to identify new business opportunities. But no sooner had they set foot in field, they were struck not only by the richness and complexity of such ordinary activities to the point they asked themselves if these practices were even measurable? Scaling from ethnography to quantitative research was not as seamless as they expected, they had to find their way to deal with two sets of data that belong to different scales if not ontological worlds. Are these two scales really strictly separated? Can't there be a way to combine them and to make them coincide? Based on the study of DIYing practices, this case study presents an attempt to integrate ethnographic and quantitative research and the challenge of resolving the scale differences between two methodologies. From turning DIYers into numbers and vice-versa, it explores the implications of ethnography, questionnaire design and data analysis.*

*Keywords: Ethnography, Mixed Methods, Statistics, Data Science, Scaling Research*

## INTRODUCTION

For quite a while, from the 19th century up to the mid-20th, the housing market was dominated by lumber dealers and served only building professionals. During the 1920s, this model started to be put under pressure. Types of building materials progressively replaced wood, homeownership increased and the market for mail-order expanded. Therefore, a new interest in home-improvement arose, as well as the products and services related to it. In response, some building-suppliers reconsidered their business models and introduced one-stop home-improvement stores catering to homeowners (Harris 2009a, 2012).

This model became progressively dominant as the Depression-Era in the 1930s and the WWII context spawned an increase in home improvement activity. Manufacturer of building material heavily supported this trend by promoting DIY with advertising, informational campaigns and by providing homeowners with credit to finance the purchases of building materials (Harris 2009b). It was in their interests after all. The fate of the industry was sealed in North America when social movements – such as "Build your own home" – addressed the post-WWII housing shortage by providing retailers stronger incentives to develop their distribution network (Harris 2012). In France, in addition to the technological change in building materials, conditions conducive to the development of the emerging home-improvement market gradually emerged. The encouragement of mortgage credit by the State in the 1930s and the post-war creation of social security enabled many solvent employees to become homeowners (Frouard 2012). Both in North America and Europe, one-stop shops customized for DIYers were now abounding on cities' outskirts.

The model hasn't changed much since then and started to decline when new players enter the home-improvement game. With the development of digital and marketplace

economy, historical retailers face a major competitors outbreak, threatening their business model. From jobbers startups to GAFA, the industry value chain seems under attack on multiple points. Putting big metal boxes on cities' outskirts, storing building materials in it and calling it a shop doesn't attract customers anymore. Traditional retailers market shares were falling, customers were leaving.

In this context of wild competition for the favors of amateurs DIYers, a major French home-improvement franchise, ask the _unknowns research & design team to help them define a new strategy. They were not just losing market shares, they were losing the meaning of their job: were they only screwdrivers and jigsaw sellers?

For the home-improvement retailer, not only was it a necessity to set a new strategy – as they were losing market shares they needed to find new opportunities and sources of profit – defining new services beyond standard retail raised a business issue as well as an HR one. As they are selling technical building materials, they need skilled and knowledgeable in-store employees to advise customers. Such employees are hard to find, hard to keep and as the digitization of the home-improvement retail is proceeding apace, their future role remains unclear. Defining a new strategy meant tackling several challenges at the same time: building new value propositions to attract customers and redefining in-store employees' role. Ideally, creating new services would serve both purposes. It would answer unmet needs and introduce a shift in employees' job, from advice and sale to something-not-very-precise-else.

Whatever the outcome could be, these challenges required a deep understanding of their customers – and the retailer had to admit they didn't' know them very much. Categorizing customers based on their purchasing history and patterns was all they had in store. So the team had to start from the very beginning and understand what DIY is and what it means in people's lives. Only after that, the team could bring a fresh perspective on their client, set a strategy and design useful services.

## DESIGNING THE STUDY

Designing new services for the home-improvement retailer meant translating their business challenges into research challenges. Since there were several unknowns in the equation, the researchers had to be methodical and thorough. They didn't know anything about home-improvement. So they had to uncover patterns as well as the social and cultural aspects behind such an ordinary practice. This is where ethnography plays its role. But it was not enough. They needed to scale up and assess to what extent these practices were spread in the general population. Were they marginal or widespread? Ethnography would only give intuitions but not a precise magnitude. The team had to turn toward quantitative techniques to assess where the market opportunities were located. In addition, as the results had to be quickly actionable, the home-improvement retailer would need clear and easy-to-collect criteria to segment the DIYers. And because the mission had to be carried out within a given budget and set time frame (of course), the methodology had to take project constraints into account. It had already been set that the quantitative part would be administered through an online questionnaire – which would leave little room for maneuver.

Ultimately, a two-pronged approach was designed:

1. Phase I: an ethnographic study that aimed to uncover home-improvement practices,
2. Phase II: a survey questionnaire administered to a representative sample of the general population that aimed to quantify the practices and DIYers characteristics identified during the ethnographic study.

## Research Strategy

Given the diversity and the complexity of building materials, tools, methods and home-improvement practices, rushing headlong into the field would have resulted in leaving the team stuck, confused by the flood of technical terms and gestures. Even more, trying to make a Prévert-ish list of all home-improvement activities could have led the team into a trap. Indeed, such a list is impossible to define a priori and impossible to make in practice: a new activity could always be added to the list. More than that, setting such a list would have led the team to consider DIY in a strict and rigid way, potentially far from the DIYers' own definition of DIY. The team tilted towards an extensive and phenomenological approach: investigate what is DIY from DIYers' point of view and what makes them engage in doing something by themselves, in their home.

## Building the Research Framework

To understand the logic of DYIng practices, the team had to take a step back to grasp DIYers and home-improvement in a broader perspective. The study was framed around several main aspects:

- The social anchors of DIY. The team's goal was to study DIY in its environment and social contexts. Who are the DIYers in terms of social origins or positions, family structures, political and moral beliefs and inclinations? The hypothesis was that home-improvement are socially situated practices (Bourdieu 1979, Bonnette-Lucat 1991).
- The embedding of DIY. The team needed to study the social structures in which DIY takes places: the family organization with its different roles, division of labor, rules, issues, way of life and how it affects DIY itself. As most activities are embedded in wider social structures (Granovetter 1985), the hypothesis was that DIY practices are shaped by the social structure where it takes place.
- The learning logics of DIY. An emphasis was placed on observing how the skills were learned in practice. From an hand-objects systems perspective (Sigaud 2012), a special focus was put on how learnings are incorporated through practice (Wacquant 2015) considering the body both as a tool for learning and action.

The main research techniques were in-depth interviews with DIYers, combined with in-home observations of improvements and alterations already realized. The team stayed for about a day at our informants home, the interviews were carried out with the DIYers and their family. A formal interview was carried out then followed by informal discussions during the visit of their home. Given the subject and the mission's constraint, the team

estimated that meeting with 15 DIYers would be sufficient enough. The guidelines for conducting the interviews were to make an archeology of DIYing: starting the conversation with DIYers as a biography through the prism of DIY and let it slip to memorable past achievements, ongoing works and future projects. For each case, the researchers focused on action-in-the-making: how the DIYers did it by themselves. Some even went so far as to play the game of "recreating" some of their achievements. It helped to grasp the practice with details and to re-stimulate the memory of informants.

Secondly, a quantitative study was carried out using an online questionnaire administered to a representative sample (N = 1200) of the French population. The original intention was to measure the behaviours and properties of DIYers and then to extrapolate the data to the general population. The initial analysis was planned as part of descriptive and inferential statistics. The sampling and constitution of the panel was carried out by a specialised market research company. The questionnaire design and the data analysis were carried out by the _unknowns team. On paper, the quantitative part seemed relatively simple, as its articulation with the ethnographic study sounded logical at least initially. It was naive, it turned out to be a bit more complicated.

Having built their research protocol, the researchers were set to go. And no sooner had they set foot in field, they understood that DIY couldn't be reduced to a predefined and limited set of activities.

## ENCOUNTER WITH THE DIYERS: THE ETHNOGRAPHIC FIELDWORK

### DIYing, Defining what DIY Is

Starting from the DIYers point of view, the researchers quickly realized that DIYing couldn't be defined so easily. Indeed, far from being boiled down to a predetermined set of activities, defining what DIY is a part of the activity of DIYing. Each DIYer builds its own definition of DIYing both in terms of activities (such as electricity, building work, carpentry or woodworking, etc.) and the intensity with which they are practised. This tension between versatility and specialization (Bonnette-Lucat 1991) is coupled with a tension between what can be done by yourself and what can (or should) be delegated to a professional.

As a result, there is no such thing as degree 0 of DIYing. Everybody is compelled to DIY to a certain extent. It would be shameful not to change a lightbulb or build a furniture kit by one's self. This social constraint implies taking care of a certain number of maintenance tasks by yourself and is rooted in family and intimate contexts and issues. Indeed, home-improvement practices are directly intertwined to a household's lifestyle and ambitions. The stakes are high: the possibility for a child to play again with an accidentally broken toy, the possibility of regaining the use of the only shower in the dwelling before a week of work. All home-improvements or alterations are made in the name of a desired way of life, couple and family agreements, and constraints of (working) life. For the researchers, the word "ethnography" took on its full meaning. From the Greek "ethnos" for family, tribe, culture, ethnographying DIYers meant studying their family culture as a whole and not DIY as a disembodied practice, floating in the air like an ectoplasm.

However, DIYing practices are fragile and constantly re-assessed through action-in-the-making. At any time, the possibility of failure can arise, which calls into question the decision to DIY.

## DIYing, Coping with Doubts and Fear

Uncertainty is a major component of DIYing: at any time, anything can go wrong. The DIYers' daily challenge is to manage these doubts and find the boldness and courage of taking action. To get a feel for the DIYers challenge, one of them, Michael a 35 y.o. engineer living in Paris recalled the story of an epic fail, best known as the shower tray incident: *"The shower tray is made of solid stone, 115 kg [...] we laid it, we glued it, I started to lay the tiles at the same time. And during the night, I said to myself: "Uh uh, I forgot to do the levels". The next morning, I arrive, I pour water: it was stagnant..."* In the end, Mike had it changed by a professional.

The shower tray incident underlines the importance of trials and errors. Making mistakes and learning from them is an important part of DIYing. Advanced DIYers are used to make mistakes and to find ways to overcome them. Indeed, most of the knowledge and skills are built and incorporated while practicing. Having experience, using their senses and relying on them to "see", to "know", and of course to "feel" while in action, is at the heart of DIY. There is no substitute for experience, not even a good handbook. And getting experience requires trying, i.e. self-confidence, a sense of authority beyond intellectual and manual skills. Thus, the ability to deal with doubt, fear and incidents is what most differentiates DIYers as well as their proclivity to engage in diverse activities and to learn by themselves.

## 4 DIYers Profiles Uncovered

The research team observed four different DIY profiles, based on their attitudes dealing with doubts and fear, playing around with tools, and building materials.

1. **Compelled DIYers**. They only tend to do small maintenance work by themselves such as minor repairs, building furniture kit, etc. Their home-improvement practices are driven by social constraints. Some even have chosen their house because no work needed to be done. Michel, a bachelor in Clermont-Ferrand said to the team: *"the apartment was in this state, I didn't redo the wallpaper, it was generally quite clean, so it was a stroke of luck! [...] I do the minimum, I know how to drive a nail to put a frame, but I have no interest".* What prevent them to engage in DIYing was they don't have enough resources (skills, sense of authority, incorporated knowledge) to take action and overcome their fear and feeling of incompetence.

2. **Hobbyist DIYers.** They carry out maintenance and repair tasks to comply with their obligations, but they also invest in some specialized activities as a hobby. Not all DIY activities are eligible to become a hobby, this mainly concerns the creation or repair of furniture, gardening, car repairs, and all sorts of activities that don't impact on the family lifestyle nor the usage of an important home feature (such as the shower). In some of these activities, DIYers can even compete with professional craftspeople.

3. **Self-Sufficient DIYers.** They tend to great versatility and expertise. They engage in DIY activities with an ideological dimension: DIY is a valorization of autonomy, resourcefulness and individual responsibility. They value the fact that they are not dependent on anyone, sometimes to the point of challenging professionals. Mark, a 38 y.o. living with his wife and a 2 y.o. baby told the team: *"I've laid steel-pan roofing, I've done roofing before. I could even do whole roofs. Honestly, [...] if I had a job that would allow me to take a year off, you know, to be out of work... if I had a job that would allow me to do that, I could build my house by myself".* In addition, they do not hesitate to tackle home-improvement issues that could impact the family lifestyle.

4. **Semi-professional DIYers.** This profile is a continuation of the self-sufficient DIYers to the point where it may be considered turning DIY into a profession: putting themselves at the service of something other than a housing ambition, as Franck, a 43 y.o. engineer turned entrepreneur: *"The apartment I buy it partly to rent it... that is the difference, it is important. Typically, I don't have nice furniture, I bought furniture from Emmaüs [local charity] [...] this summer, we removed some cables to be able to insulate the oven [...] and put a specific circuit breaker for the plates, in 32A, because they were in 20A [...] I brought the electricity up to standard, because that was dangerous, and in terms of... When you're renting, you'd better... make sure everything is up to standard as much as possible".*

The purchased-patterns home-improvement retailer customer segmentation was surely challenged. But was it enough to make it shift? Seeing these results, the home-improvement retailers executives empathize with different types of DIYers and their everyday challenges, some even recognize themselves among the profiles presented. Ethnography surely demonstrated its interest to reveal the logics behind the practices, but a concern quickly arose. Could these findings be quantified? Indeed, even if these insights could fuel the design process to imagine new services, it was not enough to identify where the business opportunities were located – if it was in their interest in specifically targeting one or more of the DIYers profiles. The strategy couldn't be defined yet, we were still in the middle of the fold. This is where the numbers come in.

## HOW DIYERS CAN BE TURNED INTO NUMBERS: SCALING THE RESEARCH

The quantitative part was where the plot thickened. Measuring DIYing behaviours turned out to be slightly more complicated than initially expected. Indeed, it implied matching two scales of analysis of a different nature.

### Two Scales of a Different Nature

When the team started to design the questionnaire at the beginning of the quantitative phase, several problems arose. Indeed, ethnographic materials and quantitative-questionnaire data are not easy to connect. Each one has its own scale, properties and captures different aspects of reality.

As ethnography captures various kinds of traces – audio, transcription, pictures, diagrams, drawings, gestures but also field notes that reflect the lay of the land – its research unit is multifaceted and diffuse. Indeed, the ethnography grasped the complex DIYers'

experience intertwined within the symbolic and material reality of their home and family. DIYers were not alone, they had spouses or wife, children and all the family participated in home-improvement to a certain point. The ethnography captured not only the DIYers experience of DIY but a part of their tribe's experience in situations.

On its side, quantitative-questionnaire captures standardized answers to a predefined set of questions but provides little information on context – especially when administered online. Thus, its research unit is unitary and discontinuous: a series of context-blind individuals reactions to the questionnaire. The research team was aware that an online questionnaire would only provide artifacts generated by the questionnaire itself.

Do ethnography and quantitative questionnaires belong to two different and irreducible knowledge scales? Realizing these differences, it seemed to the team that no direct and continuous link could be established between these two scales. The researchers wonder if they could really quantify their ethnographic finding with an online questionnaire. They felt they were trying to fit a square peg in a round hole. If it was the case, it would have meant that it was impossible to quantify ethnographic findings. And the team would have been in a good mess! But there is only one real world, and multiple ways of describing and understanding it. Ethnography and quantitative questionnaires can only extract traces from this real world – different kind of traces. And the diffuse nature of ethnographic traces should be approximated through a standardized questionnaire. The researchers were facing the same doubt the DIYers had to go through. They decided to give it a try and started thinking by putting themselves in their informants shoes: what would they answer if they were confronted with this questionnaire?

## Designing the Questionnaire: Translating Ethnography

The behaviors observed by the ethnography could not be directly measured by a questionnaire. But a questionnaire could capture specific and revealing information about behaviors or attitudes observed in the field. In a word, the questionnaire could provide approximations and hints that would need to be interpreted and combined to make sense of its data. The researchers thus designed each question – or group of questions – as a test to provide specific information on DIY practices. And in order not to be completely off the mark, it was necessary to start from the concrete reality of DIYers, i.e. ethnography.

Thus, ethnography was a valuable resource for questionnaire design. The team focused on translating the ethnographic findings into testable hypotheses within the questionnaire. The idea was to identify items or groups of items whose responses would most differentiate DIYers – well, according to the ethnographic study. To verify the existence and relevance of the profiles identified, the team gradually designed specific tests to assess DIYers profiles, their attitudes when facing doubt, their practices or their arbitrations. To achieve phrasing relevant questions required to look for details in the ethnographic material that would have remained left out otherwise. The team needed to plunge back into the interviews to analyse in detail all the activities (electricity, plumbery, building work, carpentry and woodworking, etc.), their underlying culture, vocabulary, tasks, tools and gestures.

Getting into the DIYers' shoes: that was the team's approach. 1) Use the lessons of the ethnographic study to make hypotheses (how to identify this profile knowing this?), 2) translate them into sufficiently sensitive and specific tests, and 3) use the richness and depth of the ethnographic material collected to find the most precise and relevant items and

formulations. The researchers felt there was no such thing as a one-best-way to design the questionnaire. They tested it themselves and had it tested by other members of the unknowns team in order to assess whether the questions were unequivocally understood and to get an idea of what the results could be. It was more a case for trials and errors than a deterministic science: they DIYed.

The questionnaire was administered to a representative sample of the French population. 1,200 responses were collected in the end. At this point, the researchers didn't really know what to expect, they were not at the end of their surprises.

## HOW NUMBERS CAN BE TURNED INTO DIYERS: JUGGLING WITH SCALES IN PRACTICE

When the results came back, the team started to perform an exploratory data analysis using dimensionality reduction and clustering techniques in order to verify – among other things – if the DIYers profiles identified during the ethnographic fieldwork could be found in the quantitative data.

Spoiler alert: the team faced unexpected results. On the one hand they didn't find the exact same profiles as identified in the ethnographic fieldwork, on the other hand data analysis also led to some good surprises.

### Looking for the DIYers

While filling the questionnaire, respondents were asked to report for about 40 DIY activities – of different types, issues, and levels of difficulty – whether or not they had been carried out in their dwelling and by whom (by themselves or delegated to someone else: partner, relative or professional). The underlying hypothesis came from what the ethnographic study had revealed: that Compelled DIYers only did maintenance tasks, that Hobbyists DIYers invested in a few activities without stakes, that Self-sufficient DIYers were not afraid to engage in many types of work by themselves even with high stakes. Finally, the Semi-professional DIYers tackled all types of tasks, no matter how difficult they were. Thus, each type of DIYers should have had a particular profile of response.

To cross validate their results, the researchers ran two types of clustering techniques – Kmeans and Hierarchical Classification on Principal Components (HCPC) – after a dimensionality reduction – Multiple Component Analysis (MCA). The algorithms were running, drum roll… and three groups came out on the first draft. The team felt disappointed, if not frightened. Things seemed to go wrong.

Before plunging into the intricacies of data analysis, side note: for those who aren't familiar with Euclidean distances, minimizing inertia methods and other mathematical matters HCPC iteratively builds a hierarchical tree (the so called dendrogram) by coupling two by two the closest measures in pairs and then repeating the operation, this time coupling the closest pairs. At the end, it results in a tree whose different branches partition the data into several groups. HCPC algorithms usually suggest the smallest number of ramifications to take into account by isolating the branches with the smallest variance.

The researchers analyzed the hierarchical tree provided by the HCPC and there was no reason whatsoever to think there were four clear distinct groups. In fact it was even questionable to be able to clearly distinguish groups.
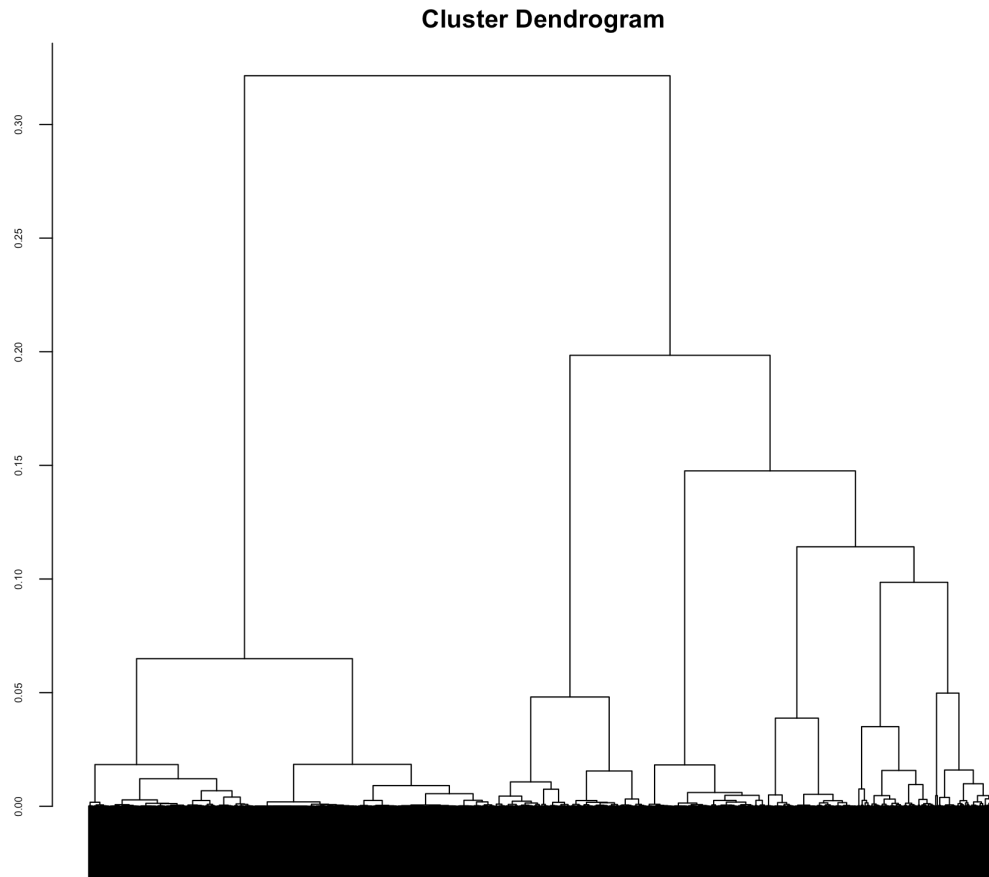
**Cluster Dendrogram**



Figure 1. Hierarchical tree built by the HCPC algorithm ©Guillaume Montagu

So the algorithm suggested three groups. Even if the researchers found that disappointing, they decided to give a chance to the machine. They looked into the data in detail and the results appeared to be even worse than expected: the groupings made by the algorithm were unclear and remained difficult to interpret. The first group seemed consistent and could match the Compelled DIYers profile previously identified, but the two others seemed to be muddled and heterogenous. That's how the researchers got a feel for the DIYers' experience with uncertainty.
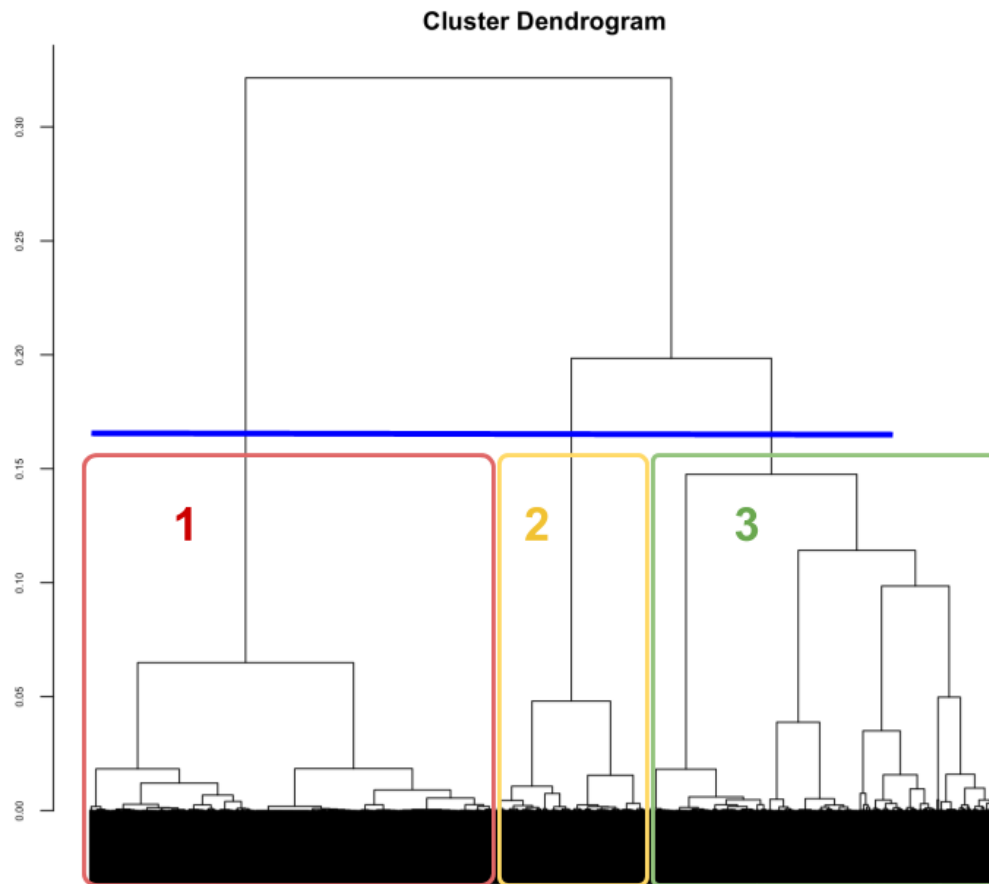
Figure 2. First hierarchical tree and cut section suggested by the algorithm that separate the tree in three groups
©Guillaume Montagu

Experiencing this dead-end, the team had to find a way to get out of this impasse. In an act of despair they decided to consider lower branches of the three. They looked at a finer level of details and considered four groups. But still, the possible interpretations of the groupings remained unclear. Without really believing in it, the team continued to descend the tree. It was only by considering nine groups that the results became clearer. The groups composed by the algorithm made more sense and, fortunately, the team found the characteristics identified in the ethnographic part.
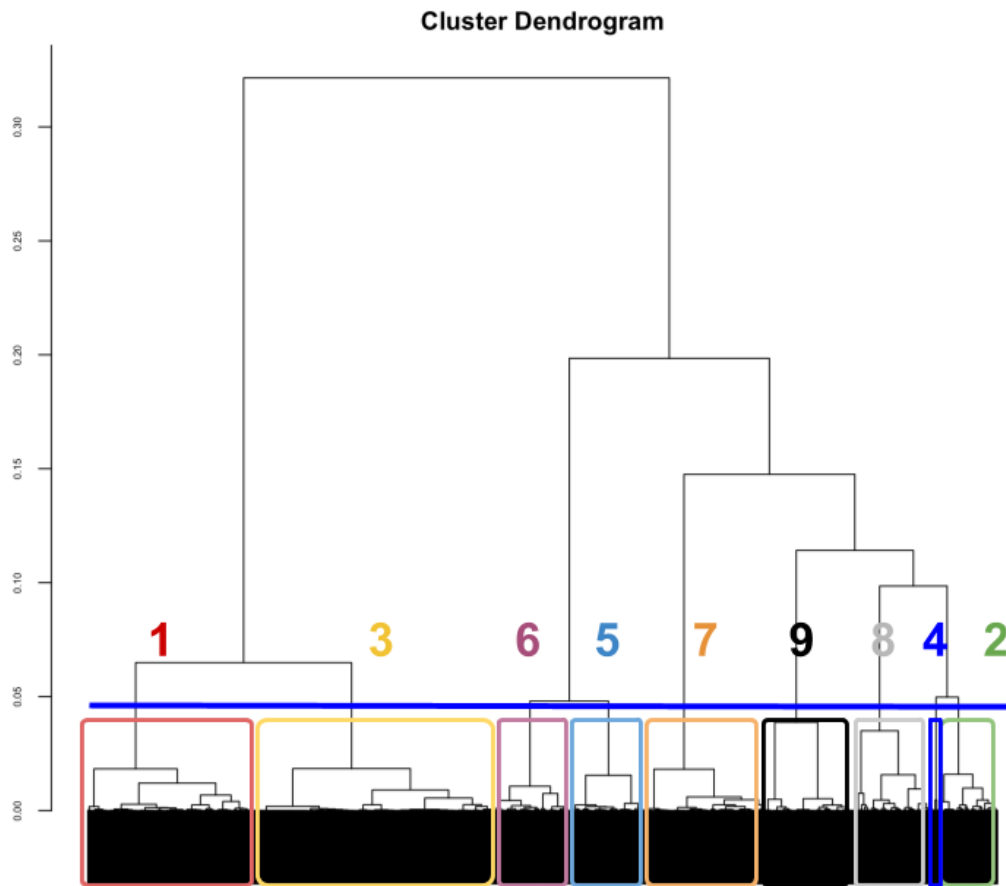
Figure 3. The same hierarchical tree with the cut section selected by the research team, it divides the data in nine groups ©Guillaume Montagu

- Groups 1, 2, 3 & 4 looked like the profile of Compelled DIYers: they stick to maintenance tasks (with some little variations). They seemed to easily give up on DIYing or getting work done. What differentiated the groups are the kind of activities delegated to professionals – usually "big" works (group 2 & 4).

- Groups 5, 6 & 9 looked like the profile of Hobbyist DIYers: beyond maintenance tasks some activities were particularly invested in, especially gardening, auto repairing, decorating, furniture repairing/building – i.e. activities that don't impact the family everyday lifestyle. Other "big" works were usually delegated to professionals. One interesting thing: group 5 looked more like they were spouses of Hobbyist DIYers, the activities beyond maintenance tasks were invested by their partner.

- Groups 7 & 8 looked like the profile of Self-sufficient DIYers: most activities were invested in whatever the complexity or importance, including technical activities such as electricity, plumbery, roofing or levelling. Another interesting thing: group 7

DIYing Along With DIYers – Montagu

looked like they were spouses of Self-Sufficient DIYers. Most of the activities and especially the "big" and technical ones were invested by their partner.

- But no clear traces of the Semi-professional DIYers at this point… Some converging hints make the team think that they were hiding in groups 7 & 8 but without enough evidence to isolate them as a distinct group from the Self-sufficient DIYers.

Why did the algorithm fail to suggest meaningful groups? The problem relied on a tacit assumption during the questionnaire design. Designing the questionnaire led the team to normalize behaviours into comparable variables. Thus, all activities were considered equivalent and comparable on the same basis. And that's exactly what the algorithm did: it attributed the same "weight" to all the variables – which is disputable. How can carpentry activities be compared on the same basis with plumbery or leveling ones? The variable couldn't be "weighted" properly to be read without bias by the algorithm. Then all the suggestions and predictions based on the inertia-minimizing criterion were not completely reliable. The algorithm did well in gathering the answers that looked alike, but did not properly identify the boundaries between groups. That was a task that could only belong to a human eye.

## Retracing the Tribes

For a moment the researchers forgot that they were interrogating individuals and not tribes anymore – as they were doing in the ethnographic part. Indeed, a "tribe" answering a questionnaire does not make sense but a member of a tribe does. And it was obvious that different kinds of "tribe" members would answer the questionnaire, especially when the sample had been drawn from the general population. Here again, the team was not operating on the same scale that what was identified first in the ethnographic material: DIYers profiles were DIYer-tribes profiles instead.

Seeing things from this perspective allowed the team to "see" the tribes in the data, beyond individual answers. It brought out a mosaic aspect of the reality of DIYers and their families. The DIYers tribes couldn't be quantified "as tribes", it was only possible to collect data from their members. Thus, the team had to reconcile this data in order to interpret it. Thus, the DIYers-tribe clusters were made on this basis and deviated from what the HCPC algorithm originally suggested.

However, the classification algorithm made it possible to highlight aspects that remained unnoticed during the ethnographic study. In particular, it was able to highlight a clear dividing line between the Compelled DIYers and the other profiles. Putting aside maintenance tasks, it became obvious and objectified that Hobbyist DIYers and Self-sufficient DIYers do more work in general in their dwelling than compelled DIYers – by themselves or by delegating it to professionals. Seeing that Hobbyist DIYers were getting significantly more work done in their homes, a new hypothesis emerged: practice begets practice, the more you DIY, the more you transform your home, by yourself or not.

But confirming and consolidating the findings of the ethnographic study was not the only goal of the quantitative part. Since the analysis was performed on a representative sample of the population, it was possible to infer the likely shares in the general population and thus the market size represented by each profile.

## WHERE KNOWING DIYERS INSPIRES STRATEGY & DESIGN

In the end, the team was not only able to draw a new customer segmentation and a strategy, they did even more. They progressively built a common knowledge, shared with the French home- improvement retailer, along this journey through different scales of analysis.

When the results of the ethnographic survey were shared, it was like an epiphany for the client. Some of the sponsors projected themselves into the results to the point of recognizing themselves in the profiles identified. It helped them to empathize with DIYers as well as understand the logic behind the practices. They knew the DIYers better and it also created a common will to challenge their existing metrics and ways to consider their customers. And this combination of scales shed a light on different aspects of their customers that they weren't aware of.

Once this knowledge base was shared, several strategic scenarios came naturally with the home-improvement retailer ending up with a reappraisal of their role in DIYers life. As a new customer segmentation arose, the question was which profiles should be targeted and how could they be targeted. This new knowledge base has irrigated all strategic work from there: on the offer, the partnerships, the distribution network, marketing, the store concept and so on.

But there was one part on which the team continued to work on. A part of the new strategy specifically addressed the Compelled DIYers. The idea was to help them to take action and launch their projects (on their own or by delegating them to a professional). The team continued the work to design a service to tackle this challenge, with the retailer employees. The knowledge built during the mission continued to be shared within the organization. A proof of concept was designed and tested at small scale in real life in five stores. Progressively, the new customer segmentation infused the home-improvement retailer culture, processes and helped them to deliver a meaningful service. As long as DIYers have to face doubts and uncertainty, there will be a need for empathy and support beyond cheaper one-day delivered building materials.

## NOTES

1. In 2012, more than 60% of French people own their own home, compared to 35% in 1954.

2. Multiple Component Analysis (MCA) was used as dimensionality reduction technique and then Hierarchical Clustering on Principal Components (HCPC) was used and crossed validated with Kmeans. All the data analysis was carried out with R and the package FactoMineR.

## REFERENCES

Becker H. S., 1998, Tricks of the Trade: How to Think about Your Research While You're Doing It, Chicago, Ill., University of Chicago Press

Benson, M., & Hamiduddin, I., 2017, Self-build homes: Social values and the lived experience of housing in practice. In Benson M. & Hamiduddin I. (Eds.), Self-Build Homes: Social Discourse, Experiences and Directions (pp. 1-14). London: UCL Press

Bonnette-Lucat C., 1991, "Les bricoleurs : entre polyvalence et spécialisation", Sociétés Contemporaines, n°8, 1991, p. 61-85

Bourdieu P., 1979, La distinction, critique sociale du jugement, Paris, Minuit.

Clogg C., 1992, "The Impact of Sociological Methodology on Statistical Methodology", Statistical Science, vol. 7, n° 2, pp. 183-207

Daynes S. & Williams T., 2018, On Ethnography. Cambridge (UK), Polity Press

Desrosières A., 2001, "Entre réalisme métrologique et conventions d'équivalence : les ambiguïtés de la sociologie quantitative", Genèses, n°43, p. 112-127

Frouard, H., 2012, "Tous propriétaires ? Les débuts de l'accession sociale à la propriété", Le Mouvement Social, 239(2), 113-128. doi:10.3917/lms.239.0113.

Glaser B. G. & Strauss A. L., 1973 The Discovery of Grounded Theory: Strategies for Qualitative Research, Chicago, Ill., Eldin

Granovetter M., 1985, Economic Action and Social Structure: The Problem of Embeddedness, American Journal of Sociology 1985 91:3, 481-510.

Harris R., 2009a, "The Birth of the North American Home Improvement Store, 1905-1929", Enterprise and Society 10,4, 687-728, doi: 10.1093/es/khp017

Harris R., 2009b, "The birth of the housing consumer in the United States, 1918–1960", International Journal of Consumer Studies, 33: 525-532, doi:10.1111/j.1470-6431.2009.00797.x

Harris R., 2012, Building a Market: The Rise of the Home Improvement Industry, 1914-1960, Chicago, University of Chicago Press,

Ladner S., 2014, Practical ethnography : a guide to doing ethnography in the private sector, Walnut Creek, CA, Left Coast Press

Ladner S., 2019, Mixed Methods: A short guide to applied mixed methods research

Sampson O., 2018, "Just Add Water: Lessons Learned from Mixing Data Science and Design Research Methods to Improve Customer Service", Ethnographic Praxis in Industry Conference Proceedings, (1):646-662, doi: 10.1111/1559-8918.2018.01227

Sigaut F., 2012, Comment Homo devint faber. Comment l'outil fit l'homme, Paris, CNRS Éditions,

Stevens, S. S., 1946, "On the Theory of Scales of Measurement", Science, 103, doi:10.1126/science.103.2684.677

Tilche A. and Simpson E., 2017, "On trusting ethnography: serendipity and the reflexive return to the fields of Gujarat", Journal of the Royal Anthropological Institute, 23, 4, p. 690-708 doi: 10.1111/1467-9655.12695

Wacquant L., 2015, For a Sociology of Flesh and Blood. Qual Sociol 38, 1–11, https://doi.org/10.1007/s11133-014-9291-y

CATALYST

# There's No Playbook for Praxis

## Translating Scholarship into Action to Build a More Ethical Bank

JEFFREY GREGER, *Varo*

*The US banking industry has a long history of excluding, exploiting, or simply ignoring low-income communities, recent immigrants, and racial minorities. In this paper, I share my experiences creating a community of practice where employees of a rapidly-growing banking startup can identify and confront the ethical challenges facing the financial technology (fintech) industry. This community is informed by insights from four years of activism and anthropological research that I conducted with small teams of service designers and ethnographers developing financial services for and with low- to moderate-income communities around the world. Through this research, I identified three institutional logics—insularity, decontextualization, and technological hubris—which limit efforts to build a more inclusive, equitable banking system. These logics hold the potential to lead well-intentioned organizations, and the practitioners they employ, to harm the marginalized communities they set out to help. This paper concludes with a reflection on the crucial role of ethnographers in identifying and counteracting ethical challenges in our organizations and industries.*

## INTRODUCTION

This year, 2020, is a unique, humbling time to be a researcher studying how people in the US manage their finances. Many people I have interviewed over the past few months are attempting to keep afloat in the midst of the second great recession of the twenty-first century. This recession has depleted savings and made it more difficult for families to make ends meet between paychecks—and, increasingly commonly—unemployment checks. To further strain household budgets, many of the largest American banks are extracting billions in overdraft and monthly account fees from the low- and moderate-income families who can least afford them, fees which wealthier customers will never see (Smith, Babar, and Borné 2020). And this only applies to people who are able to open a bank account in the first place. For those who lack the good credit scores or untarnished banking histories needed to open an account, the best option is often to use a high-fee check casher or high-interest payday lender to meet their financial needs. What I describe here is only the most recent iteration of a US banking industry that, throughout the country's history, has repeatedly excluded, exploited, or simply ignored low-income communities, recent immigrants, and racial minorities.

In recent years, neobanks (sometimes referred to as challenger banks) have launched in the US and around the world. Mission-driven neobanks promise to break with the exclusionary legacy of incumbent financial institutions by offering consumers more-accessible, lower-cost, higher-tech alternatives (Bradford 2020). But, other than their good intentions and inclusive mission statements, what prevents these companies from reenacting banking's problematic history as they grow? In this paper, I share my experiences as an in-house user experience researcher at an American neobank startup. At this startup, my colleagues and I created PeopleFirst, a community of practice where employees can identify

and confront the ethical challenges facing the financial technology (fintech) industry. This community is informed by insights from four years of activism and anthropological research that I conducted with small teams of service designers and ethnographers developing financial services for and with low- to moderate-income communities around the world. Through this research, I identified three institutional logics—insularity, decontextualization, and technological hubris—which limit efforts to build a more inclusive, equitable banking system. These logics hold the potential to lead well-intentioned organizations, and the practitioners they employ, to harm the marginalized communities they set out to help.

I begin this paper by exploring in fuller detail the history of financial exclusion in the United States. I go on to describe how these three ethics-constraining logics manifested in my research with teams of financial inclusion practitioners, and how the PeopleFirst community attempts to address these logics from within a rapidly-growing fintech company. I conclude with a reflection on the crucial role of ethnographers in identifying and counteracting ethical challenges in our organizations and industries.

## FINANCIAL EXCLUSION IN THE UNITED STATES, PAST AND PRESENT

Before delving into the logics which perpetuate financial exclusion, it is important to consider what is at stake when banks engage in unethical, exclusionary behavior, and how the financial industry and governments have responded to these challenges. The effects of unethical banking practices are etched into the American landscape. They are particularly visible in the persistently segregated communities created by redlining, the common mid-twentieth-century practice of banks refusing to lend to Black and Latinx families if they attempted to buy a home in a predominantly white community (Rothstein 2017). While redlining has been illegal since the late 1960s, its effects are still apparent in the wealth gap between Black and white families and the persistence of racially segregated neighborhoods demarcated on redlining maps. Disinvestment in these neighborhoods has also led to lower life expectancies for residents and made them susceptible to health conditions which place them at greater risk of dying from COVID-19 (Richardson et al. 2020). It was reverse redlining in these segregated communities—predatory lenders offering high-interest subprime loans to credit-starved, low-income borrowers, driven by the banking industry's appetite for these risky, yet highly-profitable loans—that was a major contributing factor in the financial crisis of 2007-2008 (Rugh and Massey 2010; Tett 2009).

The financial crisis of 2007-2008, and the subsequent Great Recession, also drew attention to a broken social contract between private banks and their public mandate to function as extensions of the Federal Reserve, a public-private partnership meant to serve the credit and deposit needs of Americans across the wealth spectrum (Baradaran 2015). From my current vantage point as a user experience researcher at a banking startup which serves many people living paycheck-to-paycheck, I have had regular reminders that many low- to moderate-income families across the US never truly recovered from the Great Recession's financial shocks, which has left them even more vulnerable in the current economic downturn. The financial crisis of 2007-2008 was but one of the many panics and expropriations of wealth which can be attributed to the banking industry, and without a deep consideration of the misaligned incentive models and ethical fault lines in the organizational cultures of banks, these harms will likely continue to happen, with disproportionate effects

on low-income communities. Throughout the history of the United States, banking institutions have been chartered with philanthropic missions to serve the financial needs of small depositors and recent immigrants (Baradaran 2015). Many of these institutions drifted toward increased profit-seeking, enabled by lax regulation, often leading banks to go bankrupt or shift toward serving wealthier depositors, to the exclusion of those living on more modest budgets.

Despite the industry's checkered past, banking remains an essential service. For those without good credit scores or access to affordable bank accounts, it can be difficult to fully participate in American society, whether a person wants to buy a home, apply for a job, or save for the future. Bank accounts can also provide quick access to paychecks, tax refunds, unemployment benefits, or disaster relief payments. For families who have little-to-no budgetary cushion between essential expenses and incomes, the timing of these deposits can determine whether they are able to purchase groceries or keep a roof over their head. Neobanks attempt to chart a new course for the banking industry. By eschewing physical branches and offering services through smartphone apps, these banks are able to offer low-fee or no-fee access to checking accounts, savings accounts, and ATMs, while providing affordable access to credit. Neobanks and other fintechs are one facet of a global financial inclusion movement which originated in international development contexts (Schwittay 2014b, 29). Financial inclusion projects often rely on efficiencies offered by mobile phones and digital infrastructures to deliver financial services that are more accessible, affordable, and responsive to the needs of low-income and rural communities who have traditionally lacked access to these services. Serving the financial needs of previously underserved communities simultaneously represents a poverty alleviation strategy and a business opportunity, resulting in frequent partnerships between governments, philanthropic foundations, and private financial and telecommunications companies.

## IDENTIFYING THE CHALLENGES AHEAD

Before starting a career in banking, I wanted to prepare myself for the ethical and organizational challenges I might face. To this end, I chose to research and think with teams of ethnographers and service designers working on financial inclusion projects. This research, conducted for my Applied Anthropology master's thesis at San José State University, centered on FAIR Money (Greger 2019).[1] This small pro bono collective of ethnographers from academia and industry came together in 2012 to research how low- and moderate-income families make ends meet in an increasingly expensive San Francisco Bay Area, with the aim of developing ethical alternatives to predatory payday loans. I was an active member of this group prior to and during my graduate research. In addition to conducting interviews with five FAIR Money colleagues (a majority of the group's active members) and documenting the group's monthly meetings, I also studied a similar project undertaken by a small design consultancy in London called Plot, which provided a comparative perspective. Early findings from my research with these two teams can be found in my 2017 EPIC paper titled "Doing Good is Hard: Ethics, Activism, and Social Impact Design as Seen from the Grassroots Perspective" (Greger 2017). In an effort to contextualize the small-scale projects of Plot and FAIR Money in the broader field of social impact design for financial inclusion, I interviewed design practitioners working on more formal projects directed and funded by large financial institutions, philanthropic

organizations, and city governments. This research also drew upon ethnographic accounts from anthropologists studying or supporting poverty alleviation projects undertaken by international aid agencies.

## THE LOGICS OF FINANCIAL INCLUSION

Through my research with teams of financial inclusion practitioners, I identified a set of institutional logics which limited the ability of these actors to address financial exclusion:

- The first logic is insularity, or an inability to learn from the ethical mistakes and near misses of others, stemming from the closed nature of many international development organizations and fintech companies.
- The second logic is decontextualization, which prevents organizations from recognizing both past and present forms of financial and social exclusion.
- The final logic is technological hubris, or placing too much faith in technologies as a cure for entrenched social inequalities.

In the next section, I discuss how these logics manifested in my graduate research with ethnographers and service designers working on financial inclusion projects as external consultants and activists. I go on to describe how these actors attempted to resist or overcome these logics.

## EXPLORING THE LOGICS OF FINANCIAL EXCLUSION

### Logic 1: Insularity

Despite the openness implied by the term "financial inclusion," the design consultancies, philanthropic foundations, and corporations involved in financial inclusion can be surprisingly insular. There are significant barriers to the free flow of information both within and between these organizations, making it difficult to learn from the successes and failures of others. Development anthropologist David Mosse (2008, 123) offers a warning to anthropologists "studying up" in development organizations who might threaten to destabilize the carefully-constructed organizational narratives involved in poverty alleviation projects: colleagues may interpret "description as an evaluation." Unauthorized attempts to describe social relations in an organization and paint portraits of individual bureaucrats can be perceived as a threat to expertise or reputations. In tech companies and development organizations, ethnography tends to be most valued when constrained to studying and making technical the user or poor "other," but it comes to be seen as dangerous when the same ethnographic lens is turned toward internal debates and decision-making processes, even if such reflexivity could further project or organizational goals. Mosse (2008, 125) explains that, given the strict "border controls" that development organizations build around their internal workings, the delicate task before anthropologists of development is to maintain their critical distance as researchers, then find ways to reenter organizations and constructively bring their insights back into the practice of development without dulling their critical edge.

The logic of insularity was also apparent in my research with financial inclusion practitioners. Many of my interlocutors were quick to acknowledge that the best way to address poverty and inequality sustainably is to put more money in the hands of low-income people, and that financial inclusion appears to be less transformative than many funders hoped it would be. Design research findings like these can be uncomfortable, unexpected, or politically contentious for clients and other stakeholders, possibly resulting in conflicts that could threaten the implementation of proposed design solutions. Conversations about organizational blind spots or contradictions may be common in academic settings, but when working with businesses as an independent consultant, these challenging conversations could lead to a practitioner being fired from a project. Consultants can find themselves in a particularly precarious position. For them, influencing the direction of a project can involve a delicate balance between having honest conversations about poverty or financial inclusion, and maintaining amicable client relationships or positive public narratives about a project's impact. The objective failure or success of a financial inclusion project to achieve its goals can become less important than legitimizing the work being done by the organization (Gould 2014, 275). Reflecting the delicacy of their positions, some of the designers and researchers I spoke with wanted to confirm the anonymity of their statements with me before offering criticisms of the financial inclusion industry or making pronouncements about the need for wealth transfers to people living in poverty.

## Logic 2: Decontextualization

Given the insular tendencies of financial inclusion projects, it can be difficult to discuss how these projects are embedded within a larger context of historical and present-day inequalities. The anthropology of international development literature provides a helpful term to think with here: the "anti-politics machine." Anthropologist James Ferguson (1994) introduced the concept of the anti-politics machine to describe the process by which the narrow bureaucratic gaze of international development institutions elides essential political aspects of poverty alleviation projects, choosing to focus instead on issues of making, implementing, and evaluating plans rooted in Western neoliberal ideology. In this depoliticized development discourse, root causes of poverty—often originating from historical disparities in power—go overlooked.

As Serena Natile (2020, 2) explains in her writing about gender discrimination related to low-income Kenyan women's use of the mobile money service M-Pesa, when new, purportedly inclusive, financial services are built atop "structural inequalities determined by intersecting relations of power, gender, race, class and poverty," they tend to reproduce these inequalities. Bill Maurer (2010, 13), anthropologist and director of the Institute for Money, Technology and Financial Inclusion (IMTFI) at the University of California, Irvine, warns practitioners and researchers involved in financial inclusion to be conscious of the place they occupy "in a long history of powerful others descending upon 'the poor' and 'their money.'" Intervention into the monetary practices of people experiencing poverty has long been a process of social control and a method for identifying and circumscribing "the poor" as a distinct target for scrutiny, control, and charitable programming. Such interventions are common in the United States. Martha Poon and Helaine Olen (2015) assert that financial literacy programs, a common project under the umbrella of financial inclusion, draw attention away from structural causes of household financial precarity, instead

promulgating neoliberal notions of individual responsibility and empowerment through programs administered by charitable organizations, government agencies, or the public education system. Personalized financial advice that helps a person navigate a tough situation can be useful, but pointing to bad money management practices as the central problem of poverty—as opposed to financial deregulation, racial wealth disparities, or stagnant wages—shifts blame onto people experiencing poverty. This opens the door to marketers from large banks and other financial experts with a "vested interest in what knowledge students do and do not acquire, encouraging certain behaviors over others, and inculcating subservience to their supposed greater knowledge and authority" (Poon and Olen 2015, 281). FAIR Money (2015) members roundly critiqued financial literacy programs for similar reasons in their report *Good with Money: Getting by in Silicon Valley*. In this report, they describe the innovative, complex strategies employed by ten people living on modest incomes to make ends meet in an expensive region. These strategies bear little resemblance to the advice given in mainstream financial literacy classes, and financial education advocates often fail to place the financial struggles of low- and moderate-income American households in the context of the country's growing economic inequality.

## Logic 3: Technological Hubris

As noted by social impact design critics such as Anke Schwittay (2014a), beginning projects with faulty preconceptions can decontextualize and individualize the financial challenges experienced by a given population, leading to ineffective or unsustainable solutions. An unfortunate effect of the depoliticization of financial inclusion projects is the belief that deep-seated social inequalities can have simple, technological solutions. Designers, policymakers, planners, and engineers have a history of unilateral action and over-simplifying or eliding political and cultural considerations when attempting to intervene in complex social systems. Evidence of this can be seen in the failures of mid-twentieth-century high-modernist utopian planned cities that presumed, and tried to enforce, ways of living based more on industrial systems and aesthetics than on how people actually wanted to live their lives (Scott 1998). Furthermore, framing a project as an endeavor to include "unbanked" people in a globalized financial system also implies a problematic hierarchy that ignores or minimizes existing networks and strategies that households already employ to cope with financial precarity.

Thankfully, ethnographic perspectives can lead organizations away from these harmful tendencies. When consulting internationally on financial inclusion projects, my interlocutor William had to fight constantly against tendencies in the technology and financial industries to characterize people as "users" whose behavior is meant to be modified by design interventions, whether to encourage consumption or for a supposed social benefit.[2] William sought to convey to his clients the agency and ingenuity he saw in the low-income communities he was researching. FAIR Money's positive characterization of payday loan recipients as "good with money" followed a similar line of reasoning. Like many other designers in this study, William's conception of good design was a process of making organizations serve the needs of their low-income clients and conform to existing cultural practices, rather than forcing people to adapt to products and services that primarily serve organizational priorities.

> People take this sort of perspective that poor people are poor because it's some sort of pathological shortcoming or character flaw, they don't work hard enough, or they're lazy, they're careless with their money. Blah blah blah blah. And it's completely ignoring the macroeconomic context, the reality of stigmatization, the legacy of slavery and Jim Crow in the United States, all these different things that are actually driving it. And so, with this work in financial inclusion internationally, I think that one of the things that we're trying to do as an organization, alongside some of our clients, is to really make that line clear, when is this an economic issue, and when is this a financial management issue.

As William describes above, an essential part of his international development practice is to identify early on which of the factors that impact a person's financial resilience are structural, and which can be addressed with a financial service. In his experience, and that of other practitioners I spoke with, financial inclusion clients tend to begin projects overestimating the value of introducing a new service or technology. Furthermore, these clients have often identified a solution before the financial problems being faced are fully understood.

## THE LIMITATIONS OF STUDYING EXTERNAL ACTORS

There were significant limitations to my graduate research, primarily due to its focus on design researchers and strategists, many of whom contributed to social impact projects as consultants. In these roles, my interlocutors were not able to offer long-term perspectives on the internal workings of their client and partner organizations. They were also primarily involved at the beginning of projects. Their contributions were often research insights and design proposals, and they rarely had the opportunity to see their ideas through to implementation, let alone post-implementation evaluation. However, this research did provide a glimpse of the challenges that lay ahead in my career in banking. It also helped me to exercise the muscle of "crossing and recrossing […] the boundary between the insider operational and the outsider researcher positions," before becoming fully immersed in the exigencies, practices, and institutional cultures which can overwhelm and blind even the best-intentioned of do-gooders (Mosse 2008, 125).

### Taking Action: Building a community of practice at work

I finally made the transition from graduate student and external actor to financial industry insider during the spring of 2019, when I joined the user experience research team at a rapidly growing fintech company. This US-based neobank startup is led by a self-proclaimed "reformed banker" and a cadre of tech and financial industry veterans with a mission to provide customers—many of whom live paycheck-to-paycheck—with an alternative to the punishing fees and outdated technologies they would be subjected to as customers of many traditional banks. The company offers a fully digital banking experience, including free checking and savings accounts with no overdraft fees. As they attempt to address financial exclusion through better technology, fintechs like this neobank confront a unique combination of the emergent ethical challenges facing Silicon Valley technology companies and the exclusionary legacy of the banking sector. During my first year working inside this neobank, I found opportunities to test, translate, and act upon the findings of my

previous research with external teams of financial inclusion practitioners in an effort to help this company build a more inclusive kind of bank.

In this section, I discuss my experiences working with a small team of coworkers to facilitate the creation of third spaces away from the demands of the bank's daily activities, in which employees could discuss the ethical challenges they encounter in their work and consider the company's role in society. A key step in this journey was identifying and working with unexpected allies from all corners of the organization—from marketing to anti-money-laundering to human resources—in order to launch an ethics and social responsibility working group within the company. I go on to discuss how this group began to address financial inclusion logics as they manifest inside the Silicon Valley fintech industry.

## Learning from "Ethics Owners"

My research with financial inclusion designers and researchers provided external perspectives on the discourses and incentive models that could lead neobanks and other fintechs engaged in the project of financial inclusion away from their world-changing aims. Coming across an article by Jacob Metcalf, Emanuel Moss, and danah boyd (2019) late in 2019 completed the picture by helping me to understand how these challenges manifest inside fintech organizations, as encountered by in-house ethical change agents. Their article, "Owning Ethics: Corporate Logics, Silicon Valley, and the Institutionalization of Ethics," contains a detailed account of the pitfalls and logics which make it difficult for established Silicon Valley companies to recognize and sustainably address ethical concerns. The authors of this article interviewed a group of 17 "ethics owners," employees who had been tasked by their organizations in recent years to drive policies and practices meant to mitigate the tech industry's well-publicized potential for harm. The logics that Metcalf, Moss, and boyd (2019) identify—which they term market fundamentalism, meritocracy, and technological solutionism—resonate with the institutional logics identified in my research with financial inclusion practitioners, which I further explore later in this section.

Frustratingly, Metcalf, Moss, and boyd (2019, 474) do not offer concrete steps to push back against industry logics that constrain efforts to address ethical challenges. Instead, they conclude their article with the following pessimistic outlook for ethics owners:

> If ethics is simply absorbed within the logics of market fundamentalism, meritocracy, and technological solutionism, it is unlikely that the tech sector will be able to offer a meaningful response to the desire for a more just and values-driven tech ecosystem.

However, they do offer a provocation to find better ways of "doing ethics" in order to bring about a "more open, just, and critical everyday practice." In light of the seemingly inescapable logics surrounding financial inclusion, what role could I play in helping to prevent the ethical pitfalls that I had spent so many years studying, once I finally set foot inside the fintech industry? I just needed to remember that industries and corporations are made up of people. Where there are people, there are potential allies and opportunities for cultural change (Heyman 2004, 491). I drew upon the final, crucial finding of my work with financial inclusion practitioners: communities of practice that bridge disciplinary and

organizational boundaries are the key to keeping conversations of ethics and social responsibility alive.

## Stepping Outside the Logics through Communities of Practice and Third Spaces

My involvement with ethnographic communities of practice began when FAIR Money welcomed me to their table (literally and figuratively). Together, over monthly potluck meals, we learned about the financial challenges facing low- to moderate-income Bay Area households, and the strategies they employ to get by in one of America's most expensive regions. I continued to witness the power of communities of practice throughout graduate school as a regular attendee of Ethnobreakfast meetings.[3] These monthly breakfasts bring together Bay Area ethnographers and "ethno-curious" folk—ranging from students to experienced professionals—interested in discussing issues related to ethnographic research in industry contexts. Ethnobreakfasts are typically hosted by a user research team in their company's office. The hosts pick a discussion topic, often looking to think with the community about a current project or to share approaches for effectively socializing research insights within an organization. As one might imagine, there is significant overlap between the Ethnobreakfast and FAIR Money communities. These communities of practice thrive in (and are co-constitutive of) third spaces, spaces like the FAIR Money and Ethnobreakfast tables where ethnographers can step away from the pressures of their day jobs to engage in conversations about the role they play as citizens of the broader world, as well as to consider the ethical implications of the emergent technologies they help to shape.

The term "third space" is typically associated with physical public or semi-public meeting spaces (e.g., libraries and coffee shops) or, more recently, virtual forums (Graham and Wright 2015). These are spaces away from work or home where people informally gather and often engage in political discussion and action. But, as Ethnobreakfast demonstrates, there are opportunities to open third spaces within the corporate context. By creating ethics-focused communities of practice inside the fintech industry, it becomes possible to address the key conflict that Metcalf, Moss, and boyd (2019, 450) identify: ethics owners attempt to engage with ethical challenges stemming from tech industry logics, despite being "fully embedded within those logics." Creating an informal third space inside a neobank could provide opportunities for colleagues to disembed themselves from fintech industry logics and address the ethical challenges of financial inclusion with clear eyes. Two years after writing about the importance of third spaces in the conclusion of my 2017 EPIC paper, I found myself inside a banking startup, with an opportunity to build such a space (Greger 2017, 289).

## INTRODUCING: PEOPLEFIRST

This opportunity came in the form of my bank's fall 2019 hackathon. Instead of designing an innovate new feature for our app, I worked with a small team of allies to launch PeopleFirst, an employee-led ethics and social responsibility working group. We created PeopleFirst to help articulate the bank's social mission and incubate policies, practices, and a culture that will keep us accountable to our mission. A core feature of this group is its monthly, seminar-style gatherings, usually over lunch, which are open to participants from

across the company. These events often begin with a volunteer giving a short presentation on a topic about which they are passionate; then the floor is opened for a guided discussion where participants can identify specific recommendations to leadership and propose topics for upcoming meetings. This format, borrowed from Ethnobreakfast, provides a space for colleagues to step back from the all-consuming work of building a bank, where we can discuss ethical issues in fintech, and consider how our company can play a positive role in society. We keep these conversations going through PeopleFirst's dedicated Slack channel, where we can discuss articles and people can propose topics for future monthly gatherings. As I describe next, PeopleFirst provides a platform from which to resist and overcome the logics of insularity, decontextualization, and technological hubris as they manifest within the fintech industry, logics that might otherwise lead a well-intentioned fintech to stray from its inclusive mission.

## Logic 1: Insularity and the Need to Open Up

Even more so than the non-profits and development institutions discussed earlier, fintechs can have well-policed boundaries within and without. Proprietary knowledge and non-disclosure agreements can make it so that companies only learn from their competitors' most well-publicized ethical missteps. Within large organizations, siloing and competition for resources can prevent teams from having honest, cross-disciplinary discussions about why a potentially harmful product might have been scrapped before launch. This leads to what Metcalf, Moss, and boyd (2019, 459) describe as the ethical pitfall of "blinkered isomorphism" at work inside Silicon Valley companies and industries. If the internal deliberations that lead a company to cancel a potentially unethical product are never publicly discussed, it is possible a competitor working along similar lines will reach different conclusions and go ahead with launching a harmful product. These well-policed organizational boundaries are reminiscent of those encountered by participants in my graduate research. Tightly controlled public (and internally directed) narratives, non-disclosure agreements, and other obfuscatory practices make it difficult to understand the true impacts that corporations have on the lives of the people they are meant to serve. These practices also impede critique and the meaningful participation of consumer advocates or marginalized groups in the design process. Another aspect of this isomorphism is that an organization's efforts to address ethical concerns can become more performative than substantive.

Isomorphism works alongside the logic of meritocracy and serves to weed contentious political considerations out of social responsibility efforts, both of which make it easier to pin ethical failures on individual employees while ignoring deeper-seated industry-spanning structural challenges (Metcalf, Moss, and boyd 2019). Similarly, corporate ethics initiatives are circumscribed by market logics, meaning they can be seen as valuable in mitigating reputational risks or preempting regulations, up until the point where they run counter to investor pressures or cause a company to cede market share to less-scrupulous competitors.

*The Response: Creating a Permeable Organization*

It was in an effort to push against these industry tendencies that we designed PeopleFirst as a distributed and employee-led forum where teams who rarely interact can learn from one

another about the challenges they face, and any failures or successes they have had in addressing them. Beyond this, PeopleFirst collaborates closely with employee-facing diversity and inclusion groups at this neobank, placing these important efforts on a continuum with more customer-facing, product-focused ethics and social impact conversations.

## Logic 2: Decontextualization and the Need for History

The next logic that PeopleFirst addresses is decontextualization. This involved bringing up a subject taboo in perpetually future-oriented Silicon Valley: history. While positive, forward-looking financial *inclusion* is a common topic of conversation in the fintech industry, there is often little space to talk about the forms financial *exclusion* can take. But it can remain a challenge to fit this essential context into bullet points, terse Slack messages, and a culture of TL;DR (too long; didn't read) common in many corporations. For employees of fintech companies attempting to bring about greater financial inclusion and build the future of banking, it can be difficult to keep focused on the roots of financial exclusion and exploitation in the systems they have been tasked with disrupting. Without recognizing the historical and social contexts we inhabit, we risk building new systems atop existing biases and structural inequalities without ever confronting these underlying factors.

*The Response: Finding Throughlines*

One repeated topic of PeopleFirst conversations has been racial discrimination in banking, which took on an even greater sense of urgency at our company following the summer 2020 Black Lives Matter protests. We discussed how redlining and other exclusionary banking practices during the mid-twentieth century helped to segregate neighborhoods and exclude people of color from home ownership, which in turn contributed to the racial wealth gaps we have today. Recent incidents of racial profiling in the banking industry, where tellers have suspected Black customers of fraud when they tried to deposit legitimate checks, have sparked conversations about preventing bias in risk policies and regulations (Flitter 2020). A PeopleFirst presentation by members of our artificial intelligence and machine learning (AI/ML) team illustrated how bias in the algorithms that banks use to assess a person's creditworthiness, or that Facebook uses to determine who sees a subprime credit advertisement, can lead to similar discriminatory outcomes.

## Logic 3: Moving from Technological Hubris to Relevant Interventions

Now we come to a final fintech industry logic: technological hubris. The challenge here is to recognize that the deep-seated social inequalities underpinning financial exclusion cannot always be addressed with technological solutions alone, and that new technologies often bring about their own unforeseen ethical risks. Metcalf, Moss, and boyd (2019) refer to the logics of meritocracy and technological solutionism as major impediments to the efforts of ethics owners. The contemporary financial inclusion industry draws upon a deep-seated faith in digital technologies (e.g., mobile-phone-based money transfers, money-management apps), and a belief in the ability of market actors and their public partners to effectively

leverage these technologies for the benefit of society by rapidly bringing financial stability and investment to underserved, low-income communities around the world.

The relentless hope invested in design, technology, and the power of markets to address entrenched social issues can lead to hubris. World-changing narratives are an integral component of Silicon Valley corporate cultures and marketing. While I do not dispute the profound social impacts of the ideas and technologies emanating out of Silicon Valley, the 'disruptions' they introduce into the world are not inherently beneficial to humanity. Hubris and a misplaced faith in technological progress can have catastrophic consequences for vulnerable populations—such as when the algorithms that US city and state governments use to automate the administration of social services flag symptoms of poverty as child welfare risk factors, or when these systems suddenly drop a person's medical coverage based on faulty or misinterpreted data (Eubanks 2018). This misplaced faith can also lead to more subtly pernicious outcomes, in the form of increased government or corporate surveillance and manipulation through the internet.

In the past few years, critical flaws have become glaringly apparent in Silicon Valley's digital utopias, through massive data breaches and the concerted efforts of foreign governments to influence the 2016 US presidential election. Facebook's former motto, "move fast and break things," might have been a more acceptable ethos when the company was a scrappy social media startup, but today, Facebook's speed and valorization of failure have proven to be incompatible with its responsibilities to society as a multi-billion-dollar organization that billions of people depend on for their news and entrust with their data (González 2017).

*The Response: Refocusing on Appropriate Technologies*

PeopleFirst has provided a platform for sharing a key insight from both my past research with FAIR Money and my current research with neobank customers: despite popular narratives to the contrary, people with the least money tend to be the best at day-to-day money management. After decades of stagnating wages and the increasing cost of groceries, housing and childcare, it is unreasonable to expect an innovative savings program or financial literacy tool to have much effect on a family's finances. In this instance, many of the solutions such as a basic income and higher minimum wages are political, not technical. Given the weight of the financial industry's exclusionary history, and the profound responsibility banks have as custodians of peoples' money, there is a real need for humility when developing new solutions. But humbly considering the limits of financial inclusion technologies does not stop innovation, it focuses it. Grasping the limitations that technologies have in addressing issues of inequality and poverty helps us to zero in on opportunities where technological solutions can better serve the needs of families living paycheck to paycheck. The current pandemic provided an excellent example of the positive social impacts our company's services can have. With accessible digital bank accounts, people previously shut out of the financial system could instantly receive unemployment and government stimulus checks, rather than waiting weeks or months for a paper check to arrive, an essential service at a time when families are operating on razor-thin margins. PeopleFirst is meant to be a thriving creative space for identifying opportunities to develop appropriate technologies which can support the existing informal financial innovations and social networks that low- to moderate-income households rely on to get by.

## KEEPING ETHICS EVERGREEN: THE FUTURE OF PEOPLEFIRST

In this group's first year, we have taken small but important steps to resist and overcome ethics-constraining financial inclusion logics. Our company's sudden switch to remote work due to COVID-19 put some of PeopleFirst's spring 2020 activities on hold, but the group is now looking toward the future, with a roadmap that involves exploring financial inclusion success metrics, ethics accountability mechanisms, and—based on the example of our artificial intelligence and machine learning team's presentation—having different teams around the company lead open discussions about the ethical challenges specific to their work. One of the greatest risks Metcalf, Moss, and boyd (2019, 456) identify is the tendency of companies to frame the problem of ethics itself as a technical problem amenable to technical solutions, sending ethics owners in search of the right checklist or static set of best practices to "solve" ethics for a company. Addressing emergent challenges and avoiding complacency is an unending process that relies on continuously making space for difficult conversations and building communities of practice that span organizations and industries. PeopleFirst is an evolving effort to establish one such community to help a neobank fulfill its financially inclusive mission as it grows. After these meetings, my colleagues echo a sentiment I heard during my time with FAIR Money: despite the often-depressing nature of our discussion topics, learning, acting, and building community with colleagues can be meaningful, enjoyable, and energizing. Not only can these benefits improve employee engagement, but they are also crucial to sustaining an informal effort like PeopleFirst. The ethics owners described by Metcalf, Moss, and boyd work inside large, established companies grappling with entrenched organizational cultures. PeopleFirst provides an alternative (or parallel) path to the employee activism seen across the tech sector, which is often a post-facto reckoning with decisions to take objectionable government contracts or failures to address problems with products that have already shipped (Scheiber and Conger 2020). Inside a relatively young fintech startup, cultures have yet to coalesce and there are opportunities to build counterbalances to fintech industry logics. At a PeopleFirst meeting early in fall 2020, for instance, I had the opportunity to present this paper, acquainting new colleagues with the group and the goals behind it, our accomplishments, and the challenges ahead.

## REFLECTING ON WHERE AND HOW TO ACT

Through both my research with financial inclusion practitioners and my experiences applying that research inside a banking startup, deciding where and how to pursue social change is a persistent challenge before would-be ethical change agents. Whether working at a grassroots level or practicing more formally as consultants and employees of financial institutions, being an ethnographer involved in financial inclusion means living in a state of constant tension. On the one hand, working closely with people who are harmed by extractive fees, race-based discrimination, or predatory lending can lead to a desire to act boldly in the face of injustice. On the other hand, we see how deep the social inequalities go that manifest as financial exclusion. Participants in my graduate research encountered dangerous incompatibilities between business cultures and social responsibility as they attempted to understand and intervene in the financial lives of low-income families. These

ethnographers repeatedly drew attention to the ways in which the simplifications, speed, and technological orientation of corporate design often sat uneasily alongside colonial legacies, historical inequalities, and the complexities of a person's socially embedded financial decision-making processes. We can become paralyzed with the knowledge of how easily well-intentioned actors can harm those they are trying to help when they take bold, hubristic actions while enmeshed in organizational and industry logics and discourses which perpetuate those inequalities. We tend to have difficulty moving beyond what James Ferguson (2010, 166–167) describes as "a politics largely defined by negation and disdain." Ferguson challenges change agents to replace the question "what are we against?" with a more challenging one: "[w]hat do we want?"

However, in spite of these dangers, sometimes radical social change needs the spark of a little hubris. During a panel discussion on international aid, anthropologist Carolyn Rouse explained that "[h]ubris has two sides. So, on the one hand, it's arrogant, it's racist. You think of American imperialism. On the other hand, without hubris, where would we be right now as humans?" (Princeton University 2015). Creative solutions to intractable problems often require bold—if flawed—visions that galvanize support from unlikely allies.

## THE CHALLENGE OF ACTION

Some financial inclusion practitioners I interviewed during my graduate research decided they could have the greatest impact working directly with corporations, governments, and large nonprofits, even if this work was not explicitly associated with making a "social impact." Those working with and within these larger organizations found themselves constantly needing to assert their expertise and the value of human- and community-centered perspectives in the design process, while subtly attempting to counteract the dehumanizing and depoliticizing effects of technocratic and neoliberal discourses. These designers and researchers needed to find ways to respectfully present and institutionalize often-uncomfortable ethnographic research insights within their client organizations while avoiding alienating project stakeholders or future employers. Others I encountered during my research chose to act externally. Although many FAIR Money members had day jobs working for and attempting to do good within Silicon Valley technology companies, the broad scope and polemical, politicized nature of their financial inclusion activities led them to a more autonomous and informal form of external activism.

When deciding how to act, the choice seems to be to either make small, yet tangible, changes in the lives of low- to moderate-income families by researching and developing services that counteract financial exclusion's manifestations, or to work to combat industry and organizational logics which cause and perpetuate financial exclusion. With PeopleFirst, we are attempting to do both. In some ways, this group could be considered hubristic. The idea that this group could overcome the logics of financial exclusion from within the fintech industry and somehow avoid cooptation may be naïve. But, as one FAIR Money member told me, "doing nothing is morally reprehensible."

Whether working as citizen activists, academics, employees, or consultants, ethnographers have the unique expertise to examine, challenge, and transform organizations of which we are a part. I take my inspiration from generations of anthropologists, from Franz Boas to Margaret Mead to Sol Tax, who have argued that ethnographers have the ethical obligation to use their positions and knowledge for the betterment of humanity.

Corporate contexts constantly challenge the ethnographers who work within them as they attempt to pursue these political commitments. Avoiding cooptation by industry logics, or paralysis when attempting to translate insights into action, requires that we continuously, publicly ask ourselves, in the words of corporate anthropologist Melissa Cefkin (2009, 18), "What are we doing there?" Wherever and however ethnographers choose to act, it is imperative that we recognize that we have the power to move beyond detachedly documenting suffering or theorizing systems of inequality. Building upon Nancy Scheper-Hughes's (1995, 420) call to "practice an anthropology-with-one's-feet-on-the-ground, a committed, grounded, even a 'barefoot' anthropology," ethnographers can become advocates and defenders of fundamental human rights, and architects of more inclusive institutions. Remaining reflexive as practitioners in industry requires the maintenance of this tight dialectical, co-constitutive relationship between social theory and political practice, a relationship known as praxis (Kozaitis 2000; Baba 2000, 33).

## CONCLUSION

I am lucky to be part of an organization receptive to a community like PeopleFirst, where I have found like-minded allies and felt welcome to bring the content of my scholarship with FAIR Money and other financial inclusion practitioners into my daily work. I recognize that, while it may not always feel like it, my user experience researcher role regularly provides opportunities to "own ethics." The longer time horizons involved in my team's work and the opportunity to interact with customers during generative user research allow us to step outside the myopia of production pipelines to consider the potential societal impacts of a product. Through interviews with our customers, we are constantly reminded of the hopes, fears, and relationships bound up in money. Through frequent partnerships with colleagues across my company, and across the fintech industry, I am able to share customer stories, reminding myself and my colleagues of the ethical stakes involved in financial inclusion. This role also allows me the privilege of shedding light on ways to better serve people who have been ignored, excluded, or harmed by banking as usual.

PeopleFirst represents a nascent effort to share this ethnographic perspective with others in my organization, and to create a space where we can articulate our values and confront the social implications of the work we do. Through PeopleFirst, we have begun to combat the fintech industry logics which underlie financial exclusion. We have done this by

- replacing organizational insularity with openness and human connection;
- introducing the historical and cultural complexity underlying financial exclusion;
- and tempering hubris with humility, by recognizing the limitations of technological solutions in addressing inherently social problems.

Writing this paper has provided a sort of third space, forcing me to step back from the daily pressures of ethnographic practice and reconnect with the scholarship, activism, and communities that led me to join an inclusion-minded neobank. I appreciate being able to share this story with EPIC, this global community of practice, in hopes of finding more allies determined to build institutions that are responsive, just, and ethical.

**Jeffrey Greger** is a UX Researcher at Varo Bank. His work focuses on the ethical and organizational challenges that design professionals face as they develop financial services for

and with low- to moderate-income communities. He holds a master's degree in Applied Anthropology from San José State University.


## NOTES

The views expressed here are those of the author and do not reflect the official policy or position of Varo Money, Inc.

Acknowledgments – I appreciate everyone who took the time to read early drafts of this paper and provide feedback and encouragement. My thanks go to the colleagues who helped launch PeopleFirst. Thanks also to FAIR Money and everyone who participated in my thesis research.

1. Portions of this paper are adapted from my master's thesis, "The Silicon Valley Approach to Poverty: Humanitarian Designers at Work in Financial Inclusion" (Greger 2019).

2. Pseudonym.

3. Bay Area ethnographers interested in attending or hosting an Ethnobrakfast can learn more at: https://sites.google.com/view/ethnobreakfast/home

## REFERENCES CITED

Baba, Marietta L. 2000. "Theories of Practice in Anthropology: A Critical Appraisal." *Annals of Anthropological Practice* 18 (1): 17–44.

Baradaran, Mehrsa. 2015. *How the Other Half Banks: Exclusion, Exploitation, and the Threat to Democracy.* Cambridge, Massachusetts: Harvard University Press.

Bradford, Terri. 2020. "Neobanks: Banks by Any Other Name?" Payments System Research Briefing. Federal Reserve Bank of Kansas City. https://www.kansascityfed.org/publications/research/rwp/psrb/articles/2020/neobanks-banks-any-other-name.

Cefkin, Melissa, ed. 2009. "Introduction: Business, Anthropology, and the Growth of Corporate Ethnography." In *Ethnography and the Corporate Encounter: Reflections on Research in and of Corporations*, 1st Edition, 1–37. New York: Berghahn Books.

Eubanks, Virginia. 2018. *Automating Inequality: How High-Tech Tools Profile, Police, and Punish the Poor.* New York, NY: St. Martin's Press.

FAIR Money. 2015. "Good with Money: Getting By in Silicon Valley." FAIR Money. https://fairnetwork.org/2015/03/16/good-with-money-2/.

Ferguson, James. 1994. "Development and Bureaucratic Power in Lesotho." *The Ecologist* 24 (5): 167–81.

———. 2010. "The Uses of Neoliberalism." *Antipode* 41 (January): 166–84.

Flitter, Emily. 2020. "'Banking While Black': How Cashing a Check Can Be a Minefield." *The New York Times*, June 18, 2020, sec. Business. https://www.nytimes.com/2020/06/18/business/banks-black-customers-racism.html.

González, Roberto J. 2017. "Hacking the Citizenry?: Personality Profiling, 'Big Data' and the Election of Donald Trump." *Anthropology Today* 33 (3): 9–12.

Gould, Jeremy. 2014. "Positionality and Scale: Methodological Issues in the Ethnography of Aid." *Occasional Paper* 0 (24): 263–90.

Graham, Todd, Daniel Jackson, and Scott Wright. 2015. "From Everyday Conversation to Political Action: Talking Austerity in Online 'Third Spaces.'" *European Journal of Communication*, December.

Greger, Jeffrey. 2017. "Doing Good Is Hard: Ethics, Activism, and Social Impact Design as Seen from the Grassroots Perspective." *Ethnographic Praxis in Industry Conference Proceedings* 2017 (1): 273–92.

———. 2019. "The Silicon Valley Approach to Poverty: Humanitarian Designers at Work in Financial Inclusion." Master's thesis, San José State University.

Heyman, Josiah McC. "The Anthropology of Power-Wielding Bureaucracies." *Human Organization* 63, no. 4 (2004): 487–500.

Kozaitis, Kathryn A. 2000. "The Rise of Anthropological Praxis." *NAPA Bulletin* 18 (1): 45– 66.

Maurer, Bill. 2010. "Monetary Ecologies and Repertoires: Research from the Institute for Money, Technology and Financial Inclusion. First Annual Report: Design Principles." School of Social Sciences, University of California, Irvine.

Metcalf, Jacob, Emanuel Moss, and danah boyd. 2019. "Owning Ethics: Corporate Logics, Silicon Valley, and the Institutionalization of Ethics." *Social Research* 86 (2): 449–76.

Mosse, David. 2008. "International Policy, Development Expertise, and Anthropology." *Focaal* 2008 (52): 119–26.

Natile, Serena. 2020. *The Exclusionary Politics of Digital Financial Inclusion: Mobile Money, Gendered Walls.* Routledge.

Poon, Martha, and Helaine Olen. 2015. "Does Literacy Improve Finance?" *Public Understanding of Science* 24 (3): 272–284.

Princeton University. 2015. "Fung Forum Panel 5: Follow the Money." YouTube video. Panel discussion at the Fung Forum, Dublin, Ireland, November 3, 2015. https://youtu.be/ceQKv8FO3Nc?t=31m49s.

Richardson, Jason, Bruce C. Mitchell, Jad Edlebi, Helen C.S. Meier, and Emily Lynch. 2020. "The Lasting Impact of Historic 'Redlining' on Neighborhood Health: Higher Prevalance of COVID-19 Risk Factors." National Community Reinvestment Coalition. https://ncrc.org/holc-health/.

Rothstein, Richard. 2017. *The Color of Law: A Forgotten History of How Our Government Segregated America.* Illustrated Edition. New York: Liveright.

Rugh, Jacob S., and Douglas S. Massey. 2010. "Racial Segregation and the American Foreclosure Crisis." *American Sociological Review* 75 (5): 629–51.

Scheiber, Noam, and Conger, Kate. 2020. "The Great Google Revolt." The New York Times Magazine, 30–34.

Scheper-Hughes, Nancy. 1995. "The Primacy of the Ethical: Propositions for a Militant Anthropology." *Current Anthropology* 36 (3): 409–40.

Schwittay, Anke F. 2014a. "Designing Development: Humanitarian Design in the Financial Inclusion Assemblage." *PoLAR: Political and Legal Anthropology Review* 37 (1): 29–47.

———. 2014b. "Making Poverty into a Financial Problem: From Global Poverty Lines to Kiva.Org." *Journal of International Development* 26 (4): 508–19.

Scott, James C. 1998. *Seeing like a State: How Certain Schemes to Improve the Human Condition Have Failed.* New Haven; London: Yale Univ. Press.

Smith, Peter, Shezal Babar, and Rebecca Borné. 2020. "Overdraft Fees: Banks Must Stop Gouging Consumers During the COVID-19 Crisis." Center for Responsible Lending. https://www.responsiblelending.org/sites/default/files/nodes/files/research-publication/crl-overdraft-covid19-jun2019.pdf.

Tett, Gillian. 2010. *Fool's Gold: The Inside Story of J.P. Morgan and How Wall St. Greed Corrupted Its Bold Dream and Created a Financial Catastrophe.* Reprint Edition. New York: Free Press.