EPIC2019

Ethnographic Praxis in Industry Conference Proceedings

Ethnographic Agency in a Data Driven World

NADINE LEVIN, Facebook

This paper argues that ethnographers can gain increased agency in data-driven corporate environments by increasing their quantitative literacy: their ability to create, understand, and strategically use quantitative data to shape organizations. Drawing on the author's experience conducting strategic user research at a technology company, the paper explores how the ability to engage with quantitative data can increase ethnographers' independence and autonomy within organizations, and can also up-level the role and value of qualitative research. The paper also explores how a deep familiarity with quantitative data can enable ethnographers to imbue quantitative data itself with new forms of agency, and can ultimately give ethnographers the tools to change institutions from within. With a greater understanding of how quantitative data is made and used, ethnographers can ensure that data is collected in representative ways, point out the limitations of existing metrics, and argue for new ways of measuring and understanding social life.

INTRODUCTION: A SOCIOCULTURAL ANTHROPOLOGIST AMONGST THE ENGINEERS

The world of modern technology companies, similar to other contemporary business environments, is full of numbers. These environments are made up of and thrive on data, which takes the form of measurements collected by apps and phones, survey results, metrics, monetary revenue, and key performance indicators (KPIs). In data-driven business environments, ethnographers often work alongside teams of engineers and data scientists who communicate with statistics, or in organizations where large-scale patterns drive decision making. Doing ethnography in these contexts means, increasingly, getting familiar with quantitative data[1] (Knox and Nafus 2018).

And yet, ethnographers often lack a fluency and familiarity with quantitative data, which can limit their engagement with numbers and the social world in which they a pervasive. In the academic world, social anthropological training—in contrast to other social science disciplines like economics or psychology—often includes little to no exposure to statistics or programming languages. In business environments, ethnographers face large barriers to learning quantitative skills like statistics and coding. Researchers in the already liminal field of user experience research must produce results on fast timelines, and may not be able to justify the time and space needed to learn new quantitative skills. As a result, ethnographers often have less familiarity with quantitative data. They may lack the skills to analyze quantitative data or write code. They may struggle to interpret quantitative data in the form of database tables, statistical models, or charts and graphs. Or, they might not have the deep understanding of data that is required to use quantitative data and insights to drive strategic organizational change.

This distance from data is further evidenced in the anthropological and social studies of science (STS) literature on quantitative data, which has tended to focus on its end products—papers and visualizations, societal implications (Miller 2015; Dougherty 2015)—instead of the everyday practices that go into negotiating and making sense of quantitative data (Levin 2014a; Starosielski 2015; Buur, Mosleh, and FYHN 2018). This has left ethnographers, in both industry and academic settings, with few playbooks for engaging critically with quantitative data, or for doing "participant observation" on the practices that shape how quantitative data are made and used.

This paper takes as its problematic the author's three years of experience working in a large, data-intensive organization, where ethnographers often struggle to gain traction with their work because of an inherent bias towards quantification and numbers (Maiers 2018). By providing an autoethnography of my experiences learning quantitative methods (Jones, Adams, and Ellis 2016), I argue that quantitative literacy—a deep familiarity with data, as well as knowledge of how to use quantitative skills to shape an organization—can give ethnographers expanded agency[2] to do research in and impact data-intensive businesses. To be effective industry researchers, ethnographers must not only deliver insights into their subject area, but must also understand and move within the larger systems in which these insights operate (Cefkin 2010). Fluency with coding and statistics can give ethnographers increased independence and authority within organizations, by enabling them to speak the same language as their stakeholders, and by helping them more clearly articulating how quantitative and qualitative research can complement each other. This ability to not only work, but also tap into the social power of numbers, becomes a powerful tool for gaining ethnographic agency.

However, quantitative literacy not only gives ethnographers new agency, but can also enable them to imbue quantitative data itself with new forms of agency and meaning. Quantitative skills can empower ethnographers to ask crucial questions about how data is used to make decisions in organizations. This can, for example, highlight crucial gaps in datasets, making space to question whether business outcomes are measured with the correct metrics. Quantitative literacy can give ethnographers the agency not just to critique institutions, but also to change them from within. With a greater understanding of how quantitative data is made and used, ethnographers can ensure that data is collected in representative ways, point out the limitations of existing metrics, and argue for new ways of measuring and understanding social life. With quantitative literacy, anthropologists can gain the tools to re-negotiate and restructure the quantitative environment around them, by changing the processes through which data shape and have power in the world.

And yet, it might not always be possible, necessary, or ideal for researchers trained in ethnography to pursue quantitative literacy. If researchers do not have the skills, opportunity, or desire to develop quantitative literacy, does their impact suffer? Does this emphasis on quantitative literacy create false expectations that everyone can and should be able to tackle both the quantitative and the qualitative? Do qualitative approaches become relegated to the less influential projects and parts of the company? Might promoting a more quantitative way of viewing users and technology make it unintentionally harder in the long run to get buy-in for qualitative approaches and insights?

Ultimately, this paper asks, how can we better equip ethnographers to have more agency as they enter this world of quantitative data, statistics, and algorithms? What are the best approaches and theories to help ethnographers work and succeed in data-intensive environments? How can the agency that comes with quantitative literacy help ethnographers have more impact, by enabling them question and influence the structures and values surrounding data-driven decision making? And ultimately, what does this need to have quantitative literacy in order to gain a "voice" within technology companies say about the current culture of user experience research, or the current culture of knowledge and values within data-driven organizations?

PART 1: GAINING AGENCY BY LEARNING THE LANGUAGE AND SKILLS OF QUANTITATIVE DATA

In this first section of the paper, I reflect on my experiences becoming fluent in creating, transforming, understanding, and using large datasets at a well-known tech company. I describe the process of carrying out a large-scale survey project, where I learned how different sampling methods impacted survey data, wrote code to analyze this data with statistics, and told stories about this data to stakeholders to influence company strategy. Through my work, I came to do participant observation with quantitative data not in a cursory, surface-level way—as can happen through reading papers or relying on second-hand accounts in interviews—but by deeply engaging in the creation, analysis, and socialization of quantitative data. Drawing on this experience, I reflect on how quantitative literacy can help ethnographers have more autonomy and influence in an organization, and can simultaneously uplift the status of qualitative research methods and insights in data-driven environments.

For my PhD, I did an ethnography of how scientists worked with large datasets and statistics. I began my research at a laboratory at Imperial College London thinking that I would study academic-industry collaborations, and ended up focusing on how scientists were trying to understand the complex system of metabolism with multivariate statistics (Levin 2014a). I carried out participant observation with scientists in the field of "metabolomics," as they attempted to understand the role that metabolism played in enabling living beings to interact with their environment over time. I watched scientists and clinicians put samples of urine, blood, and tissue into mass spectrometry and nuclear magnetic resonance machines, and also observed how they analyze the ensuing datasets—which contained hundreds of thousands of data points, and could be several gigabytes large—with statistics and algorithms.

Ultimately, these statistics and algorithms became the focus of my participant observation with scientists. Although ethnography is typically thought of as a distinctly qualitative methodology, anthropology, it turns out, has always had a relationship to numbers (Curran 2013). Adam Kuper writes, for example, that early British anthropology had an "overriding concern with the accumulation of data" (Kuper 1977, 5), and that Malinowksi looked to collect "statistical documentation through concrete evidence" as part of his ethnographies (Kuper 1977, 15).

Consequently, as I carried out participant observation with scientists who were using complex, black-boxed machine learning algorithms (Eubanks 2018), I came up with particular ways to engage with these data practices. I shadowed researchers as they did lab experiments and analyzed data in MATLAB and other statistical software. I pored over scientific literature and attended training courses and seminars. Even though I had no formal training in statistics, I learned to "speak the language of data" by familiarizing myself with the theory behind principal components analysis, supervised learning techniques, and neural networks (Levin 2014a). By gaining a deeper understanding of virtual and intangible data-rich systems, I was able to reflect on how quantitative data was reshaping concepts like metabolism and health, creating friction between scientists and clinicians (Levin 2014b), and shaping notions of "persons" and "populations" in healthcare systems.

Although I spent much of my PhD thinking and writing about how data was impacting society (Levin 2018), in academia, I did not need to learn how to actually *do* data analysis. To write an ethnography about data and statistics, it was enough to shadow scientists, to understand the theory behind statistics, and then to talk to participants about how data was impacting their understandings of metabolism and health. My success as an academic was in no way tied to my ability to do quantitative work, and as such, I had no incentive or reason to learn to do data analysis.

In my industry job, however, I found myself in an environment where having quantitative skills seemed to unlock a number of doors and opportunities. Before I began working in the tech industry, newspapers and magazines seemed to portray ethnography in business world as an almost mystical tool for unearthing consumer insights (Wood 2013; Singer 2014) or for enacting organizational change (Huhman 2018). In my own experiences, however, most of my colleagues—other user researchers included—did not understand the nature or value of qualitative data. They did not understand which business questions would best benefit from ethnographic inquiries, or that qualitative insights were never meant to be "representative" (Maiers 2018). As a result, business decisions were still largely driven by quantitative insights from surveys and by behavioral shifts seen through the lens of log data.

In my everyday world, qualitative data was often seen as a storytelling tool, and little more. "Qual," as it was colloquially called, was relegated to the "human" or "ethical" dimension of big data, rather than existing as an equal form of data in and of itself (Arora et al. 2018). I often advocated that qualitative data could be used to come up with new product directions, or to develop principles and values for product design. But because I was not producing data that would neatly fit with existing metrics—in the format of a survey that said "4% of people thought X with product Y"—many of my stakeholders did not know how to operationalize my qualitative insights.

After about a year in my job, as a mostly qualitative researcher whose interaction with data was limited to an occasional analysis of survey data in Microsoft Excel, I realized that my lack of quantitative expertise was preventing me from engaging in strategic conversations in the company. I struggled to engage with data scientists and engineers, as they talked about the numeric results of A/B experiments. I also found it challenging to engage with other user researchers who came from more quantitative social psychology backgrounds, as they talked about complex survey analysis in the programming language R, with which I had no familiarity.

I found myself caught in a double bind. I wanted to advocate for qualitative methods within the organization, but I realized that I needed to become more adept with quantitative methods to do so. On the 0ne hand, I wanted to grow my identity within the organization as an ethnographer and anthropologist, but on the other hand, I wanted to have access to new projects, relationships with stakeholders, and forms of impact. I started to wonder if a deeper understanding of the ways that quantitative data was being created, manipulated, and used would help me understand and influence the data-driven organization in which I worked.

Working in a fast-paced company, I did not have the luxury of continuing my doctoral research, by doing extensive interviews or participant observation as the primary method of becoming more fluent with quantitative data. Instead, I needed a more practical approach. I began to learn how to conduct more complicated survey projects, as a way to develop a greater familiarity and understanding of quantitative data, and also as a way to strategically advance my position within the organization. But as I started to improve my knowledge of survey design and sampling, I still had to rely on data scientists and quantitative researchers for help with querying databases, or with figuring out the right statistical tests to use on my data.

At the time, I was still using Excel to analyze my data. I had learned this computer program during my undergraduate studies, as I had taken classes in biology or chemistry that only required simple data manipulation like adding or calculating p values. But with survey data, I began running into an increasing number of problems with data manipulation, particularly as the datasets became bigger and more complex. Each time I analyzed a subset of my survey data or carried out a new type of analysis, I had to create a new tab within an Excel workbook. This led to a proliferation of tabs and hand-coded calculations, leading to issues with version control and mistakes with calculations.

Observing that the "expert" quantitative researchers throughout the company were analyzing their data in R, an open source statistical software package, I decided to follow in their footsteps. To learn R, which has a notoriously complex syntax (Machlis 2017), I took advantage of the several in-person and online training courses that my company offered. But learning R syntax abstractly, without concrete datasets to solve for, was challenging. As a result, I designed a moderately complicated survey, and started practicing data analysis with my own datasets. Knowing that I wanted to answer specific questions about the data, I was able to translate my working process in Excel into R, by looking at "R Cheat Sheets" (https://www.rstudio.com/resources/cheatsheets/), drawing out visual diagrams for how the data should be manipulated, and by debugging issues on the website Stack Exchange (https://stackexchange.com/).

As I engaged more deeply with quantitative methods than I had during my PhD, I went through a process of becoming fluent not just with the language, but also in the skills of quantitative data. Beyond a high-level, theoretical understanding of statistics, I learned how to conduct representative sampling with large surveys, how to effectively structure survey questions to control for response bias, how to join survey data to other data in our databases, how to write code to analyze data quickly and efficiently, and how to tell stories with numbers through graphs and other visuals. As I became more fluent in the skills needed to manipulate quantitative data, I also developed a greater understanding of the types of research questions that would best benefit from qualitative versus quantitative approaches, as well as how the two methods could be combined to drive the greatest impact. I could more clearly identify when research—take for example a study to understand which strategies people were using to learn new things—would benefit from a survey rather than a qualitative study.

As a culmination of my efforts to learn more about quantitative data, I completed a large survey project, which delivered a number of insights that shaped company strategy. I used stratified sampling and weighting—key strategies for minimizing bias—to illuminate the complexities of the product's user base. This caused stakeholders to question their assumptions about how and by whom the product was being used. Instead of presenting the survey results as an average, which would have lumped the experiences of different populations into one number, I showed how the survey results varied depending on where the user lived, how old they were, and whether they used an android or apple phone. The "user," which had formerly been an amorphous concept (Amirebrahimi 2016), was suddenly anchored in rich contextual information. As a result, my stakeholders were forced to consider how social environments shaped peoples' interactions with technology, because there was incontrovertible quantitative evidence that gave texture and shape to a formerly "average" user.

These kinds of insights, however, would likely not have emerged had I partnered with a quantitative researcher or data scientist instead of doing the work myself. Because I was no longer reliant on others to work with data—to query databases or carry out statistical analyses—I gained the autonomy and freedom to approach data with my own unique perspective. Once I understood the common ways that researchers approached surveys, I began to question the decisions that were made about which differences in data to highlight, or about how data should be visualized. I approached my survey data with the eyes of an ethnographer, with a view towards drawing on the multitude of dimensions in the data, in order to highlight cultural, social, and regional differences.

As my quantitative work became more visible in the organization, somewhat unexpectedly, more stakeholders began to pay attention to my ethnographic work. As I paired qualitative and quantitative approaches within larger projects, the people I worked with began to understand how in-depth interviews and ethnographic insights were both valuable types of data, which were part of a larger story that could be told about a problem space. By uncovering a number of interesting trends in the survey, I had created new opportunities to do ethnographic research with specific populations. While the quantitative survey data provided the "what," qualitative methods like ethnography helped uncover the "why," the reasons underling the differences in the data.

Doing quantitative work ultimately became a way to elevate the status of my qualitative work throughout the organization. Because my ethnographic analysis became data-driven, my stakeholders perceived it to be more rigorous and high quality. The close relationship between quantitative and qualitative research helped to circumvent the all-too-common criticism that qualitative research lacked statistical validity or situational generalizability (Maiers 2018), Here, ethnography was not just a way to give texture to quantitative data. Instead, quantitative methods emerged as a way to give new value and life to ethnography and qualitative data itself, by leveraging "big data" to open up opportunities to explain cultural differences (Curran 2013).

My experiences with quantitative research not only led to new opportunities for qualitative research, but also transformed my role and status in the research organization. Following this survey project, I was given license to do more strategic projects—and even assumed a new role as a "pathfinding" researcher, focused on the future of the business—because I had learned to deliver insights in a shape and format that the organization recognized and understood. I was able to take on projects that were bigger in scope and spanned longer timelines, as I was no longer classified as a "qualitative researcher" and could now address complex topics using whatever method I needed. As a result, researchers in the organization began to solicit my help and advice with tackling complex problems, transforming me into a trusted thought-partner for driving company strategy.

My push to create more impact by upskilling in quantitative methods did, however, have some unintentional consequences for ethnography and ethnographers within the organization. Research leadership began to promote the hybrid quantitative-ethnography approach that I had developed as an "ideal" model for other researchers. And yet, other ethnographers who had less exposure to quantitative environments or less flexibility to pursue skill development in their free time, struggled to adopt this model. By emphasizing the intertwined nature of surveys and ethnography, I had helped to create false expectations that anyone, regardless of their background or resources, could and should gain experience in quantitative methods as a pathway to having greater impact. While quantitative literacy increased exposure to and understanding of ethnography in some ways, it also devalued the method and placed it in a more precarious situation in other ways.

In summary, this section of the paper shows how by developing quantitative literacy, ethnographers can gain the ability and agency to gain more autonomy and influence in data-driven organizations. While organizational cultures and social norms—like the dominance of quantitative data and reasoning—can create power asymmetries that make it difficult for qualitative research to have impact, ethnographers are not helpless. Just as patients with mood disorders can repurpose "constraining" technologies like in-vitro fertilization or brain imaging for their own strategic ends (Lock and Kaufert 1998; Cohn 2004), ethnographers can "hijack" quantitative methods for their own strategic ends within organizations.

Here, for example, quantitative literacy can become a strategic tool to help ethnographers gain back some of the agency that was lost when ethnography was fit into the user-experience framework (Amirebrahimi 2016). When this occurred, the focus on and language of "the user" flattened research into the binary of the user and the used, removing much of the richness of peoples' local, social, and culturally-specific engagements with technology. While ethnography's multifaceted engagement with culture and power is often reduced to the individual usage of a device, quantitative data can help bring context, specificity, and place to qualitative data, by showing how technology usage varies by dimensions like age, gender, and country.

Ultimately, in "expert" environments like tech companies, ethnographers can more critically and meaningfully engage with technologies like databases and algorithms by becoming fluent in the language and end-to-end processes of data. While such expert knowledge may not be necessary in the context of the quantified-self movement (Nafus 2016), a lack of expert knowledge in tech companies can preclude ethnographers from participating fully in the social life and strategic decisions of organizations. In this way, possessing certain skillsets, or not possessing others, can alter power dynamics and disrupt the so-called "big data divide" that exists internally in organizations.

PART 2: IMBUING QUANTITATIVE DATA WITH NEW AGENCY, BY SHAPING THE NORMS, VALUES, AND POLITICS OF NUMBERS

In this second section of the paper, I reflect on the process of driving impact and decision-making with a large-scale, hybrid ethnographic and quantitative project to measure the relationship between digital skills and product usage. I talk about how *doing* quantitative work, rather than just observing it, can give ethnographers critical insight into the politics of data in a large institution: into what is and is not being measured with quantitative data (Crawford 2013, 2016), into how the contingency of data is negotiated in decision making (Latour and Woolgar 1986), and into the ways that certain forms of data come to be valued and have power (Räsänen and Nyce 2013; Biruk 2018; Rajan and Leonelli 2013).

During my PhD, I had used my understanding of the entire lifecycle of metabolic data to develop a theoretical toolkit for approaching data practices in the laboratory. By carrying out ethnography in a data-intensive environment, which might appear off-limits or intimidating to anthropologists, I came to see how numbers were not "stable and objective measures of reality" (Biruk 2018), but had complex social lives and were embroiled in power dynamics. In this way, I began to reflect on the social aspects of how data were made, reasoned through, and used, and how these "data practices" shaped how people and societies functioned. My work with numbers gave me insight into the politics of data (Boyd and Crawford 2012), and gave me a theoretical toolkit (Levin 2018) for analyzing the claims to objectivity (Daston and Galison 2007), newness (Boellstorff and Maurer 2015), and accuracy that surround quantitative data (Gitelman 2013).

As I transitioned out of academia, I was confronted with similar claims and concepts around quantitative data in an industry environment. I saw how data practices affected strategic decisions around which business needs and populations should be a priority (emerging markets or western markets?), and also around which metrics should be used to measure success (consumer satisfaction or the number of active users?). Becoming deeply involved in quantitative research was a way to apply the theories I had developed during my PhD to my industry work—theories which could ultimately help me understand and function in the organization. Although I was technically doing "user" research, quantitative methods enabled me to scrutinize not just the end users of the system, but also the system itself.

Focusing on the practices that create and shape data, as well as on the organizational structures in which data operates—in what Julia Haines refers to as "multi-dimensional ethnography" (Haines 2017)—can help overcome common dichotomies like quantitative versus qualitative research. For example, actor-network theory (Callon 1984), which sees various "actors" operating as "nodes" within a network, has successfully shown how both human and non-human entities, like machines and data, can have agency (Mol 2002). But, as Marilyn Strathern points out in *Cutting the Network*, the webs of inter-relations that connect the nodes in networks are not all evenly spaced and distributed (Strathern 1996). Networks, like numbers, have distinct qualities, such that some connections between nodes are longer or shorter than others. Agency and power can be unevenly distributed in networks, highlighting how some points of view—like quantitative insights from log data and surveys—have more power than others—like qualitative insights from ethnography.

I became fully immersed in the politics of data when I began a project to understand and measure how issues with digital literacy were leading to negative product experiences. Leading up to this project, my ethnographic work with older adults in the California Central Valley, as well as with people who were newer to the internet in Vietnam, had revealed how phone interactions that Silicon Valley often took for granted—uploading and posting pictures, formulating Google searches—were difficult for some populations. As various stakeholders at the company began to ask what role digital skills played in the amount of time or frequency that people engaged with digital products, I carried out international fieldwork in Brazil and Indonesia. The goal was to understand the range of problems people with low digital skills encountered, and to identify how these problems were different than the problems frequently assumed or encountered during research in Silicon Valley.

This fieldwork identified a number of design problems that people with lower digital skills encountered, such as not understanding how hidden press-and-hold gestures[3] worked, or not understanding how to navigate a complex product. My qualitative data also indicated that when people with low digital skills had trouble interacting with the product, they were more likely to experience problems with safety and well-being. For example, if someone did not know that privacy as a concept existed, they might be more likely to share information to a wider network than they realized, revealing personal information to strangers.

Alternatively, if someone did not understand that they could report behavior that was overtly sexual or violent, they might continually be exposed to harmful or negative content. Ethnography was crucial to generating these insights, as it gathered feedback from people who typically did not participate in surveys, or who might not have the knowledge or vocabulary to describe their problems during more cursory qualitative research.

Ultimately, this research spoke to a fundamental gap in knowledge at the company, and also in the academic and non-profit world. Studies of digital skills were almost entirely conducted in North America and Europe, leaving out the experiences of the majority of the world's population. To address these gaps, I began advocating for a set of product changes—more simple user interfaces, more education to help people understand complex concepts, spaces within the product where people could build confidence when exploring new features—that would specifically solve problems for people with low digital skills. But as I struggled to convince other people in the organization to work on these initiatives, I also began to question if the business was measuring its outcomes in a way that could incentivize, or capture the benefits of, this kind of work.

For example, one of the major ways that the company measured success was by tracking growth—the number of people using the product—and engagement—how often during a month those people used the product. These measures of success, however, could sometimes be at odds with improving peoples' understanding of features, or with reducing peoples' potential to be confused by complex features. In one paradoxical example, clarifying how a certain feature worked actually ended up decreasing engagement with the product, because people became aware that they were making mistakes with the feature, and therefore began to use the features more cautiously. In another example, product improvements that reduced the spread of negative or misleading content also decreased growth, as people had less content to engage with overall. These examples demonstrated how it was extremely challenging to developing measurement frameworks that articulated the right balance between incentivizing growth and mitigating potential risks.

During my struggles to convince people to make products better for people with low digital skills, I pragmatically realized that qualitative data would not be enough. To motivate change, I needed to come up with a framework for quantitatively measuring digital skills and their impact on product metrics like growth and engagement. Despite the existence of external literature suggesting that a large proportion of the population had low digital skills (Kankaraš et al. 2016), I was constantly asked, "Can we size this?" Without a metric, I could not concretely say that people with X level of skills used the product Y percent less, and that the company would have Z percent gain in engagement if it focused on improving skills rather than launching new complex features. I realized that stakeholders would not act on qualitative insights unless they could be corroborated by quantitative data. In a business driven by numbers, product managers and engineers needed to be able to quantify the impact of digital skills relative to other factors, to ultimately decide whether they should invest in digital skills relative to other areas of opportunity.

As a result, I expanded my project on digital skills to include a quantitative phase, where I worked with academics to develop a framework for measuring digital skills across populations. As ethnographers, when we delve into and get involved in the practices of data, we do not simply represent and repeat numbers from a business angle. Instead, we have an ethical obligation to represent the needs of users of the product, and to articulate how technology relates to peoples' social realities. With quantitative survey work, I attempted to elevate the problems and needs of group of people who were not only less engaged with the product, but whose voices had also historically been silenced because the company lacked an appropriate measurement tool. Qualitative work had highlighted this crucial gap in our metrics, which I could now fill with quantitative work.

As I developed a large-scale survey to measure digital skills, I leveraged insights from my qualitative work to develop survey questions, and also to make the survey's sampling as representative as possible. I saw how some of the past survey work at the company had unintentionally excluded the voices of people with lower digital skills, by over-focusing on populations who, due a variety of structural factors, were more likely to take surveys. I drew on my deep understanding of surveys to make the process of data collection more inclusive, oversampling people who were newer to the product and used it less. I imbued my data and the process of administering surveys with new agency, by changing which voices were represented in the data. Even if surveys could not fully reach or elevate the experiences of people with lower skills in the same way as ethnography, the changes I enacted in the survey process ensured that product decisions and changes would include the experiences of people with lower digital skills.

By using my quantitative skills to run a carefully crafted survey, I began to work with product teams to suggest how we could incorporate measurements of digital skills into their product frameworks, by showing how digital skills were correlated with metrics teams were already tracking. Some teams were highly receptive to this information. They saw it as a tool for tracking which populations were more likely to struggle with products, as well as which populations were more likely to benefit from product fixes. While teams had struggled to understand and act on ethnographic findings, as I translated my qualitative insights into a quantitative form, I suddenly presented results in a language that my stakeholders spoke.

As my quantitative research elevated an awareness of digital skills throughout the company, my work began to reshape how people used metrics in product development. Teams began to focus more on new users and people using lower end phones, placing more value on the experiences of these under-represented groups. By combining ethnographic insights with survey insights, I had found a way to pragmatically navigate and make impact in a business environment that was dominated by numbers and metrics. Ultimately, I gave new agency to quantitative data by transforming who and what it represented, and also by changing the way people used and thought about it.

However, not all teams were receptive to qualitative insights, even if they were "backed up" by quantitative data. While my project encouraged stakeholders to apply design changes and principles to reduce complex product experiences, some stakeholders felt the need to quantify complexity itself. Another group within the company began to develop a "complexity metric" that could identify the numeric complexity of a given design. However, this complexity metric was rooted in computer science and psychological approaches to complexity, which did not account for the various ways that people around the world perceived and experienced "complexity" as a concept. Moreover, this complexity metric could only identify the "what" and not the "why" of complexity, leaving stakeholders with a tool to track but not fix the underlying causes of complex product experiences. Because the complexity metric entailed a purely quantitative approach, instead of encouraging the application of qualitative insights supported by quantitative data, it gained significantly more traction among some teams and stakeholders. This signaled that, even with the increased agency and influence of quantitative literacy, furthering ethnographic approaches in data-driven organizations remained a challenge.

Despite these challenges, in this section I argue that by "getting their hands dirty" with quantitative data, ethnographers can become empowered to suggest changes to data-related methods and processes within data-driven institutions. By combining qualitative and quantitative insights, or by leverage quantitative methods to support qualitative findings, ethnographers can help institutions reflect on what is missing from existing datasets, can help teams figure out if they are collecting the right data or measuring the right, and can ultimately influence how (and what types of) data are used to drive strategy.

CONCLUSION: ETHNOGRAPHIC EMPOWERMENT IN DATA-INTENSIVE ENVIRONMENTS

This paper shows how quantitative literacy—the ability not just to produce and understand quantitative data, but also the ability to use and apply it strategically within organizations—can give ethnographers the ability not just to critique institutions, but also to change them from within. By restructuring who has access to and can generate narratives about data, quantitative literacy enables ethnographers to renegotiate and restructure power relations in data-driven environments. As ethnographers learn how to do data-related tasks like running surveys, interpreting metrics and models, and writing code, they can challenge the epistemic authority of other disciplines that typically produce and control narratives about data. In doing so, they can gain a seat at the decision-making table to discuss important issues and tradeoffs in company strategy.

Having more of a strategic voice within an organization, and restructuring international power relations, is a difficult undertaking. The reality is that unless ethnographers learn to speak the language of and deeply understand quantitative data, they will struggle to enact institutional change, and to shift the power and value afforded to qualitative research. Knowing a system, rather than just the users of that system, allows for changes at the level of values and norms rather than products. This kind of thinking can reveal how and why complex social problems cannot be addressed with new metrics or algorithms alone. If the "smartness" of AI lies, as Clare Elish writes, in its power to process patterns and numbers with statistics (Elish 2018), then anthropologists need to play a role in the creation and deployment of statistical systems. Anthropologists must widen their horizons to focus not just on users and designs, but also on the machine learning algorithms, data architectures, and institutional hierarchies that make up data-driven organizations.

Ultimately, it is possible that promoting the adoption of quantitative methods and skills amongst ethnographers will increase the precarity of ethnography as a method and approach. The goal of this article is not to argue that all ethnographers *should* gain quantitative literacy, but rather that they *could*, as an avenue towards effecting institutional change. Quantitative literacy is one of many possible avenues that ethnographers can take, as they push institutions to reflect on whether data are made and used in the best and most ethical ways.

Nadine Levin is a senior researcher at Facebook, who conducts strategic work on under-represented populations, including older adults and those with low digital skills. You can contact her at nslevin87@gmail.com with any thoughts.

NOTES

I would like to give thanks to my colleague and fellow anthro evangelist Abbas Jaffer, who read and provided comments on this paper, and has been a sounding board for these ideas throughout our time together at Facebook. I would also like to thank my husband Andrew Symington for always reading and being a tireless advocate of my work.

1. Here, I use "data" as an umbrella term encompassing a variety of numeric, quantitative inputs and outputs, ranging from the datasets on which AI are trained, to the metrics companies use to make product decisions.

2. Here, I use agency to refer to the socioculturally mediated capacity to act (Ahearn 2001).

3. Common examples of this are: (1) using the "Like" button to add heart and other reactions to Facebook posts, or (2) tapping and holding an Instagram story to pause the progression to other stories.

REFERENCES CITED

Ahearn, Laura M. 2001. "Language and Agency." Annual review of anthropology 30 (1): 109-137.

Amirebrahimi, Shaheen. 2016. "The Rise of the User and the Fall of People: Ethnographic Cooptation and a New Language of Globalization." Ethnographic Praxis in Industry Conference Proceedings.

Arora, Millie P., Mikkel Krenchel, Jacob McAuliffe, and Poornima Ramaswamy. 2018. "Contextual Analytics: Towards a Practical Integration of Human and Data Science Approaches in the Development of Algorithms." Ethnographic Praxis in Industry Conference Proceedings.

Biruk, Crystal. 2018. Cooking Data: Culture and Politics in an African Research World. Duke University Press.

Boellstorff, Tom, and Bill Maurer. 2015. Data: Now Bigger and Better! Chicago: University of Chicago Press.

Boyd, Danah, and Kate Crawford. 2012. "Critical Questions for Big Data: Provocations for a Cultural, Technological, and Scholarly Phenomenon." Information, Communication & Society 15 (5): 662-679.

Buur, Jacob, Sara Said Mosleh, and CHRISTINA FYHN. 2018. "Physicalizations of Big Data in Ethnographic Context." Ethnographic Praxis in Industry Conference Proceedings.

Callon, Michel. 1984. "Some elements of a sociology of translation: domestication of the scallops and the fishermen of St Brieuc Bay." The sociological review 32 (1_suppl): 196-233.

Cefkin, Melissa. 2010. Ethnography and the Corporate Encounter: Reflections on Research in and of Corporations. Vol. 5. Berghahn Books.

Cohn, S. 2004. "Increasing Resolution, Intensifying Ambiguity: An Ethnographic Account of Seeing Life in Brain Scans." Economy & Society 33 (1): 52-76.

Crawford, Kate. 2013. "The Hidden Biases in Big Data." HBR Blog Network. https://hbr.org/2013/04/the-hidden-biases-in-big-data.

---. 2016. "Can an Algorithm be Agonistic? Ten Scenes from Life in Calculated Publics." Science, technology & human values 14 (1).

Curran, John. 2013. "Big data or 'big ethnographic data'? Positioning big data within the ethnographic space." Ethnographic praxis in industry conference proceedings.

Daston, LJ, and P Galison. 2007. Objectivity. New York: Zone Books.

Dougherty, Conor. 2015. "Google Photos Mistakenly Labels Black People 'Gorillas'." The New York Times. http://bits.blogs.nytimes.com/2015/07/01/google-photos-mistakenly-labels-black-people-gorillas/.

Elish, Madeleine Clare. 2018. "The Stakes of Uncertainty: Developing and Integrating Machine Learning in Clinical Care." Ethnographic Praxis in Industry Conference Proceedings.

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

Gitelman, Lisa. 2013. Raw Data is an Oxymoron. Cambridge, MA: MIT Press.

Haines, Julia Katherine. 2017. "Towards Multi-Dimensional Ethnography." Ethnographic Praxis in Industry Conference Proceedings.

Huhman, Heather R. 2018. "Why Corporate Anthropology Is the Missing Link in Your Company's Culture." Inc.com. https://www.inc.com/heather-r-huhman/why-corporate-anthropology-is-missing-link-in-your-companys-culture.html.

Jones, Stacy Holman, Tony E Adams, and Carolyn Ellis. 2016. Handbook of Autoethnography. Routledge.

Kankaraš, Miloš, Guillermo Montt, Marco Paccagnella, Glenda Quintini, and William Thorn. 2016. "Skills Matter: Further Results from the Survey of Adult Skills. OECD Skills Studies." OECD Publishing.

Knox, Hannah, and Dawn Nafus. 2018. Ethnography for a Data-Saturated World. Oxford University Press.

Kuper, A. 1977. Anthropology and Anthropologists: the British School 1920–1970. Harmondsworth: Penguin.

Latour, B., and S. Woolgar. 1986. Laboratory Life: The Construction of Scientific Facts. Princeton: Princeton University Press.

Levin, Nadine. 2014a. "Multivariate Statistics and the Enactment of Biological Complexity in Metabolic Science." Social Studies of Science 44 (4): 555–578.

---. 2014b. "What's Being Translated in Translational Research? Making and Making Sense of Data between the Laboratory and the Clinic." TECNOSCIENZA: Italian Journal of Science & Technology Studies 5 (1): 91-114.

---. 2018. "Big Data and Biomedicine." In The Palgrave Handbook of Biology and Society, 663-681. Springer.

Lock, M., and P.A. Kaufert. 1998. Pragmatic Women and Body Politics. Cambridge: Cambridge University Press.

Machlis, Sharon. 2017. "Beginner's guide to R: Syntax quirks you'll want to know." https://www.computerworld.com/article/2497319/business-intelligence-beginner-s-guide-to-r-syntax-quirks-you-ll-want-to-know.html.

Maiers, Claire. 2018. "Reading the Tea Leaves: Ethnographic Prediction as Evidence." Ethnographic Praxis in Industry Conference Proceedings.

Miller, Claire Cain. 2015. "When Algorithms Discriminate." The New York TImes. http://www.nytimes.com/2015/07/10/upshot/when-algorithms-discriminate.html.

Mol, A. 2002. The Body Multiple: Artherosclerosis in Practice. Durham, NC: Duke University Press.

Nafus, Dawn. 2016. "The domestication of data: why embracing digital data means embracing bigger questions." Ethnographic Praxis in Industry Conference Proceedings.

Rajan, K. S., and S Leonelli. 2013. "Biomedical Trans-actions " Public Culture 25 (3): 463-475.

Räsänen, Minna, and James M Nyce. 2013. "The Raw is Cooked: Data in Intelligence Practice." Science, Technology & Human Values 38 (5): 655-677.

Singer, Natasha. 2014. "Intel's sharp-eyed social scientist." New York Times. https://www.nytimes.com/2014/02/16/technology/intels-sharp-eyed-social-scientist.html.

Starosielski, Nicole. 2015. The Undersea Network. Durham: Duke University Press.

Strathern, Marilyn. 1996. "Cutting the Network." The Journal of the Royal Anthropological Institute 2 (3): 517-35. http://www.jstor.org/stable/3034901.

Wood, Graeme. 2013. "Anthropology Inc.". The Atlantic. http://www.theatlantic.com/magazine/archive/2013/03/anthropology-inc/309218/.