

Increasing Perceived Agency in Human-AI Interactions

Learnings from Piloting a Voice User Interface with Drivers on Uber

JAKE SILVA, *Uber Technologies*

This case study seeks to increase understanding of how agency is fostered in human-AI interaction by providing insight from Uber's development of a conversational voice-user-interface (VUI) for its driver application. Additionally, it provides user researchers with insight on how to identify agency's importance early in the product development process and communicate it effectively to product stakeholders. First, the case reviews the literature to provide a firm theoretical basis of agency. It then describes the implementation of a novel in-car Wizard-Of-Oz study and its usefulness in identifying agency as a critical mediator of driver interaction with the VUI before software-development. Afterward, three factors which impacted driver agency and product usage are discussed -- conversational agency, use of the VUI in social contexts and perception of the VUI persona. Finally, the case describes strategies used to convince the engineering and product teams to prioritize features to increase agency. As a result, the findings led to substantive changes to the VUI to increase agency and enhance the user experience.

INTRODUCTION

Founded in 2009, Uber is an on-demand transportation platform that connects riders to available drivers through a mobile app. Today, Uber is available in more than 700 cities globally and is expanding its platform services to include food ordering and delivery, e-bikes and scooters and freight logistics. To help fulfill its goal of making transportation safe and easy for everyone, Uber is developing a conversational voice-user-interface (VUI) for its driver application to enable hands-free interactions with it. User research played a critical role in shaping the VUI's design throughout the product lifecycle. First, an exploratory in-car Wizard-of-Oz (WoZ) study was conducted prior to software development to identify what functionality drivers wanted in a full-fledged conversational VUI and to observe how they interacted with it while driving. Later, three studies in the US, India and Australia were run in which drivers used a minimal-viable product (MVP) VUI for two weeks on actual Uber trips and gave feedback on the experience through SMS messages, a survey and semi-structured interviews. Insights from the exploratory WoZ led to the identification of agency as a critical mediator of driver interaction with the VUI and a factor that could impact product usage if not supported. Subsequent real-world insights from the studies in the US, India and Australia confirmed agency's impact on the VUI user experience and usage.

The emergence of AI technology like VUIs raises important questions on how the relationship between humans and technology will change moving forward. While some technologists insist that AI will soon automating everything, it is more likely to augment human capabilities for the foreseeable future rather than replace them, creating a new paradigm of greater human-AI collaboration (Simon 2019; Sloane 2019; Manyika 2019). Consequently, agency and its role in mediating human-AI interaction will require reconfiguration. If not, designers and developers risk the disuse of their technology due to

reduced human agency. Thus, this case study seeks to increase understanding of AI's impact on agency by providing insights from the development of the Uber conversational VUI. First, the case reviews the literature to provide a firm theoretical basis of agency. It then describes the implementation of the in-car WoZ and its usefulness in identifying agency as a critical interaction mediator prior to software development. Afterward, three factors which impacted agency and product usage are discussed—conversational agency, use of the VUI in social contexts and perception of the VUI persona. Finally, the case describes how the researcher overcame challenges in convincing the engineering and product team to prioritize features to increase agency.

AGENCY IN HUMAN-AI INTERACTION

Defining Agency

Agency—defined as the feeling of control over one's actions and their consequences - is a fundamental human trait that significantly influences how humans behave and interact with one another and technology. Consider for example the seeming irrationality of adding placebo buttons to crosswalks and elevators or those who fear flying but drive every day despite the latter being a statistically much riskier endeavor—agency undeniably has a profound effect on how humans behave. And while agency is commonly considered an individual phenomenon, social cognitive theory provides a more nuanced framework which acknowledges the impact of social factors beyond the individual on agency. Specifically, the theory defines three types of agency—individual, proxy and collective (Bandura 2000). On an individual level, agency is exercised through the sensorimotor system which associates actions one takes in the environment to their causal effects. For example, think about a person opening a door—they feel the force of the rotating handle and visually see the door open. On the road, a driver moving through a curve feels the centrifugal force on their body reinforcing the outcomes of their actions. But agency is not strictly a physical phenomenon. It's also influenced by one's context information, background beliefs and social norms - a concept known as judgement of agency (Synofzik 2013). In other words, people not only experience agency physiologically but also interpret it subjectively, making it an imprecise measure of reality in which a person may feel more in control than they actually are. Beyond individual control, people exercise proxy agency by influencing key social actors in their lives, like colleagues, employers and government, who have the resources, knowledge and means to help secure desired goals. Finally, the need to form groups to achieve objectives unattainable individually means people also exercise collective agency whereby groups develop shared beliefs on their capabilities beyond the individual.

Agency and Human Interaction with Technology

Agency is fundamental to people's successful use of technology through user interfaces and belief that it serves them in achieving their goals. This is especially true in the digital world where processors, logic and complex algorithms increasingly play an intermediary role between user input and their resulting actions. Take for example sending an email to a friend. While seemingly simple, pressing send triggers an intricate orchestration of largely invisible events to get the message to the receiver. Schniederman highlights the need to

support agency in this digital space in one of his ‘Eight Golden Rules of Interface Design’ saying digital products should allow users to be the initiators of actions and give them the sense that they are in full control of events (Shneiderman 2016). Familiar interface elements like hover and focus button states, mobile-phone haptics and loading GIFs as well as the point-and-click/tap paradigm of today’s digital interfaces all serve to give users this sense of control.

And while people enjoy a high degree of control in the majority of today’s digital interactions, the increasing intelligence and pervasiveness of AI technology requires a rethinking of how human agency will evolve and be fostered in human-AI interaction. But before going further, it’s useful to discuss the current evolution of traditional automation into what many term AI. In the classic sense, automation is the use of an artificial mechanism like a machine or software to partially or completely replace the human labor needed to accomplish a task. Accordingly, everything from the appliances that free us from washing dishes and clothes by hand to the autopilots that fly airplanes can be considered automation. And while deterministic tasks like these have been automated for decades, technology has reached a transition point where it’s able to perform more complex and non-deterministic tasks. Consider call centers for example. Where phone representatives once used judgement and skill to interact with customers, AI software can now coach them in real-time giving advice on speaking pace and rapport with the customer (Garza 2019). This insertion of AI to augment human capability raises serious concerns on how representative agency might be reduced necessitating management to frame it as an improvement tool rather than a replacement one.

The above example highlights the need for designers to invent new ways to foster agency-sharing so humans and AI can achieve shared goals effectively. Examples of such strategies already exist in some familiar AI-powered technology. For example, search engine autocomplete gives users automatic, yet easily dismissible suggestions that accelerate search and refine ambiguous human intents creating a mutually beneficial shared agency (Heer 2019). Proposals to augment agency in more complex interactions include making the logic behind AI decisions interpretable to users (Holstein 2018), designing shared representations of the human-AI mental model (Heer 2019) and even allowing users to self-assemble an AI system itself (Sun 2016). The risk of not supporting a user’s agency is them disusing or misusing technology - an all-too-common outcome in conventional automation. For example, reduced agency from cockpit automation can decrease a pilot’s situational awareness (Endsley 1995) and even degrade their fine-motor flying skills (Haslbeck 2016), contributing to aviation accidents. Research on collaborative human-AI systems finds similar risks with participants in one study facing significant difficulty in building agency during joint actions with an artificial partner which hindered successful task completion (Sahai 2017). Diminished agency can also harm emotionally. A 2017 ethnographic study of unmanned drone pilots found for example that reduced agency diminished pilot stature in the eyes of colleagues, negatively impacted their own self-perception and ultimately affected their career prospects (Elish 2018). Thus, more insight is needed to understand how to foster agency in human-AI interaction to mitigate the consequences of reduced agency. Or as Applin and Fischer put it succinctly “In order for this [human-robot] cooperation to succeed, robots will need to be designed in such a way that the ability for humans to express their own agency through them is afforded” (Applin 2015).

Agency in Conversational Interfaces

Conversational interfaces have emerged as an increasingly ubiquitous means to facilitate human-AI interaction through natural-language and can be found in many popular consumer applications like Siri, Alexa and Google Assistant as well as in business applications like customer support, banking and HR. While conversational interfaces attempt to make interaction with AI easier through human-like dialogues, they often lack the cues like facial expression or body language that are so critical to facilitating effective human-human interaction. This increases the risk of reduced agency because a person has less information to know when to speak, what to speak and how to speak to their artificial conversational partner - three constraints necessary for successful human-human conversation (Gibson 2000). Numerous studies show conversational interfaces embodied with additional human-like characteristics like in virtual avatars foster more agency than text or voice alone (Appel et. al. 2012; Astrid 2010). These observations reflect individual and proxy agency which were discussed in the literature review - that is, humans need both sensory feedback and the feeling they can socially influence others to effectively interface in conversation. Given that conversational interfaces replace the human element with the artificial, designers must consider how to foster shared conversational agency in conversational interfaces. To date, little research specific to voice-only interfaces exists and thus this case seeks in-part to deepen understanding of the space.

FOUNDATIONAL WIZARD-OF-OZ RESEARCH

Research Goals

To inform the design and development of the VUI for drivers, foundational research was proposed before software development began which had three primary goals. First, the research aimed to identify the information and actions drivers desired in a full-fledged VUI to inform engineering requirements and the long-term product roadmap. Second, it sought to capture how drivers would behaviorally interact with the VUI - would they speak more slowly, loudly and clearly, interrupt voice prompts or physically lean or glance toward it - to inform the VUI's visual, sound and conversational design. Third, given that VUIs are still a developing technology, the research aimed to gauge driver tolerance for voice recognition errors to influence the design of error-handling and prevent them in the first place. Understanding the impact of the VUI on driver agency was not an explicit goal. Rather, the chosen methodology and resultant findings described later in this case revealed it as a critical mediator of driver interaction with the product.

Methodology

Given the highly interactive nature of VUIs and dynamic context of driving, a one-hour in-car wizard-of-oz (WoZ) study was chosen to gain insight into driver behavior and attitudes toward the proposed Uber VUI as naturalistically as possible. In a WoZ study, a human 'wizard' simulates the functionality of a working software artifact out of view of participants, leading them to believe they're interacting with a real product. For the WoZ, an in-car on-the-road setting was selected to better ground the research in the context of drivers

and increase the fidelity of observations. However, the in-car setting and lack of a functioning prototype in Uber's driver application required significant deviation from a typical WoZ to be successful. After evaluating multiple setups, it was determined a 2x2x1-inch nano bluetooth speaker placed inconspicuously in the driver's car and connected to a text-to-speech (TTS) wizard interface could believably simulate voice interactivity in the driver application (see Figure 1). With the interface, which was built using HTML and Javascript, the wizard could inconspicuously act as the voice and respond to driver questions and requests to the VUI. To make these interactions smooth and believable, the interface had canned responses for anticipated queries like *"what's the airport queue"* or *"contact support"*. The researcher-wizard could also initiate voice interactions with the driver by sending prompts through the interface like *"incoming message from rider, how would you like to respond?"* or *"new trip, say yes to accept"*. Furthermore, an open-text field on the interface could be used to craft custom responses to unanticipated questions or requests. To maintain the illusion of a functioning VUI while crafting custom responses, a *"one moment please"* canned utterance was played. The open-source Web Speech API TTS engine powered the wizard interface and had a default standard American-accented male voice which was used during the study.

Twelve drivers - six each in New York City and Chicago - were selected to participate in the WoZ. The two cities were chosen based on their large populations, dense urban cores and diversity in cultures. To recruit participants, a random subset of drivers in each city received an email invitation to the study with details on its purpose and timeframe. Those who expressed interest in participating then completed a brief survey to self-report gender, native language, length of time driving on the Uber platform and previous experience using voice interfaces. This information was used to select a diverse participating cohort to ensure richer insights. Selected participants were then contacted by phone to confirm their participation and instructed to meet the researcher with their vehicle at a predetermined public parking lot in their city. Participation was voluntary and drivers were compensated for their time.

To set up the WoZ in participant vehicles, the researcher sat in the front passenger seat and mounted an Uber-owned smartphone pre-installed with the driver application to the front console. To add voice interactivity, the nano-speaker was placed inconspicuously in the vehicle and connected to the wizard interface on a laptop via bluetooth. Additionally, because the driver application interface was not actually controllable via voice, the researcher needed to physically manipulate it at certain moments to move the trip flow forward e.g. swiping the start trip button. To overcome this limitation and preclude questions about the purpose of the nano-speaker, the participant was told the VUI was a prototype and not yet fully integrated into the driver application. The researcher then signed into the application with a test account so the participant could drive three simulated but realistic Uber trips around the city during the one-hour study period. Participants did not pick up any actual riders during these trips but drove each one as they typically would. Finally, a GoPro camera was mounted on the vehicle dashboard to record sessions for later analysis.

With setup complete, the researcher requested the first ride of each session using an internal trip simulator tool. After a few seconds, the request appeared on the driver application with the researcher simultaneously clicking on the wizard interface *"incoming trip request, say accept or decline"*. Drivers listened and then used voice to accept the request with most saying *"yes"* or *"accept"*. The researcher then physically tapped the navigation button on the application to move the trip flow forward and display directions to the simulated rider

pickup location. During the whole trip flow—pick-up, driving and drop-off—drivers were encouraged to think about what helpful information or actions they might ask of the VUI in addition to the de facto interactions of accepting, starting and ending a trip with voice. Some of the open-ended queries drivers asked included “*call Uber support*”, “*call my rider*” and “*take me to a busy area*”. VUI-initiated interactions like “*New message from your rider, I’m on the corner of market and second street, how would you like to respond?*” were also used to prompt drivers. Several errors were introduced during each one-hour study to gauge driver tolerance for them. For example, a driver might say: “*tell my rider I’m five minutes away*” and the researcher would subsequently introduce the error: “*Sorry, I didn’t catch that. Please say again*”. At the end of each one-hour session, the driver was debriefed on their experience with the VUI and informed of the WoZ setup.

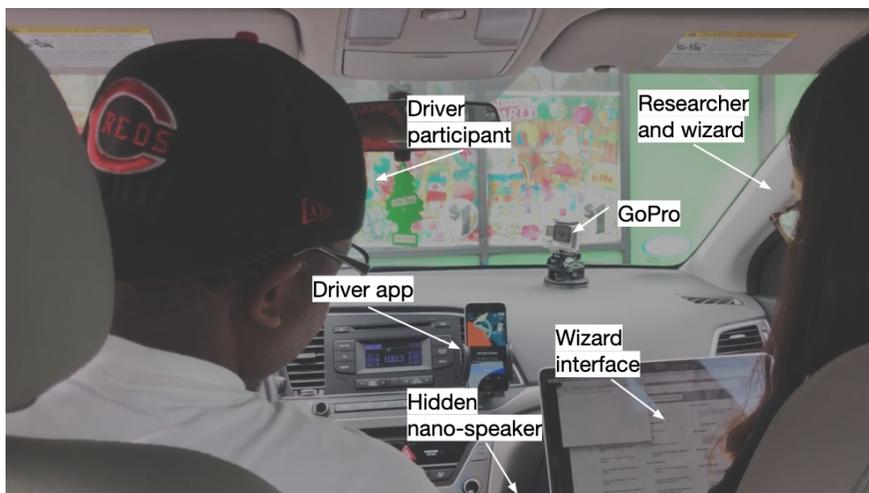


Figure 1. In-car WoZ setup.



Figure 2. Nano-speaker used to simulate voice interactivity.

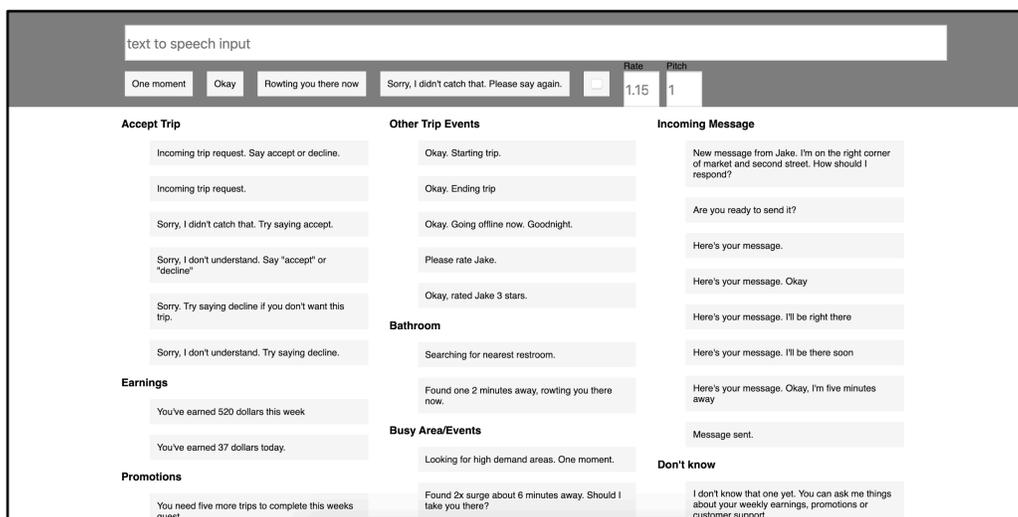


Figure 3. Wizard of Oz interface.

Findings on Agency

Insights gained from the WoZ revealed agency as a critical mediator of driver interaction with the VUI. The first piece of evidence supporting this was the observation that drivers frequently interrupted the VUI when it spoke, especially during VUI-initiated interactions. For example, the VUI would suddenly say, “*New trip request, s...*” and the driver would interrupt with “*yes*” before it could finish speaking. This ability to interrupt a VUI is known as barge-in and requires the technical ability to respond to a user request while simultaneously listening for interruptions—a complex feature to implement. Barge-in is supported in some commercial voice products like Alexa. That drivers barged-in reflects their need to have conversational agency—the feeling they can influence others in conversation to achieve their objectives, regardless of whether ‘others’ is a human or machine. Thus, fostering agency in VUIs through features like barge-in becomes critical for a good user experience. Based on this finding, it was hypothesized that launching an Uber VUI without barge-in functionality would likely reduce driver agency and consequently lead to its disuse. Accordingly, the researcher began to make the case to the engineering team to prioritize barge-in for the MVP VUI despite the significant technical investment - a process discussed later in this case.

The WoZ also identified the use of the VUI in social contexts as potentially impacting driver agency and product usage. This was learned during the session debriefs in which some drivers reported they might feel uncomfortable speaking and responding to the VUI in front of riders. While drivers would generally control when to speak to the VUI, given the way the Uber platform works, they might receive VUI-initiated prompts for trip requests or rider messages for their next trip while on the current one. The reduced willingness of some to use the VUI in this social context is likely due to the perceived stigma of talking to an artificial agent in front of other people which several studies confirm is a common phenomenon (Milanesi 2017; Moorthy 2014). It was hypothesized that if drivers did not

have the option to disable VUI-initiated interactions in this situation, it could reduce their agency and lead to disuse. However, because the insight was gained through self-reporting and no actual riders entered the vehicle during the study, it was determined the feature would remain on with riders in the car during the alpha research phase to validate the hypothesis.

ALPHA RESEARCH PHASE

Overview of the MVP VUI

The WoZ revealed agency as an important factor to consider in designing the VUI. However, to reduce time-to-market and quickly validate the value of the VUI before further development, a limited MVP was built enabling drivers to accept or decline trip requests using their voice. When a request arrives in the standard driver application, a card pops up providing key information about the trip such as time and distance to the rider. If the driver decides to accept, they tap the request card. If they want to decline, they either tap an x-icon or let the request expire by doing nothing (see Figure 4). The MVP does not remove tapping but adds voice as an additional modality to accomplish the same task. When a request arrives, the MVP prompts “*new trip, say yes to accept*” followed by an audible beep to cue the driver to speak. Drivers then say “*yes*” or other affirmative statements like “*accept*” or “*okay*” to accept and “*no*” or “*decline*” to decline. An oscillating white bar at the bottom of the request card provides visual confirmation the system is processing the utterance (see Figure 4). As previously mentioned, drivers did not have the ability to turn off the MVP in front of riders to help validate whether they would disuse it in a social context. Additionally, barge-in was not built into the MVP due to the significant technical effort required to build it. This resulted in drivers having to wait for the VUI to finish saying “*new trip, say yes to accept*” before they could respond or else the microphone would not capture their utterance. Lacking barge-in and the option to turn the VUI on or off in a social context was predicted to reduce usage of the MVP, however, the product team ultimately decided the tradeoff of building the functionality was not worth the delay in time-to-market. This is a common tension in industry between research and product in which teams often ship digital products fast to learn and iterate quickly sometimes at the cost of a perfect user experience. In this case, the researcher acknowledged the need to ship fast and validate the MVP and turned it into an opportunity to run an alpha research phase to collect real-world insights and further strengthen the case for prioritizing agency.

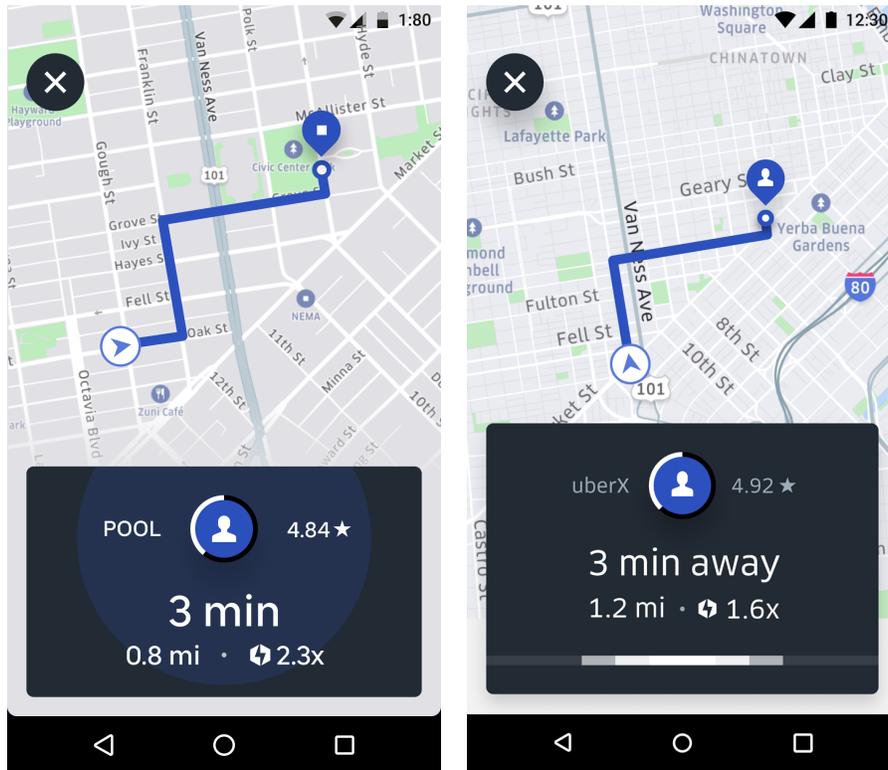


Figure 4. Standard trip request card on the left showing key trip information. To the right is the VUI request card with white listening bar at the bottom.

Methods

To gain real-world feedback on the MVP and validate the WoZ findings on agency before a wider release, an alpha research phase was initiated. Alpha research parallels the software engineering concept of the alpha release in which a rough version of software is made available to a small group of internal technical users to evaluate its stability and quality. Similarly, alpha research tests with real users within the context of a qualitative research study gaining feedback on a product's experience, not just its functionality. 38 drivers were recruited in the San Francisco Bay Area to test the MVP on real Uber trips for a two-week period. During that time, the MVP was turned on for all participating drivers. However, they were not required to use it and could still tap to accept trip requests if they desired. Providing both modalities generated real-world data that could be used to analyze agency's impact on usage - if drivers faced reduced agency in interacting with the VUI, they could tap. Similar to the WoZ, participants were diverse in native-language, gender and time on the Uber platform. A real-time SMS feedback channel, survey and semi-structured interviews were used to triangulate and strengthen findings.

Findings on Agency

Analysis of the qualitative and usage data from the alpha research confirmed that a lack of barge-in reduced driver agency during interactions with the MVP VUI and consequently their usage of the product. In some cases, drivers interrupted the voice prompt leading to their utterances not being captured. Others, who correctly waited to respond to the VUI, reported feeling stressed given the limited time to accept or decline requests. For example, one driver said in an SMS message: “[I’m] anxious...waiting for prompt so not to miss trip opportunity. Once I get a trip offer, I start saying Accept, Accept, Accept until she responds”. Those reporting this diminished agency demonstrably reduced their usage of the MVP over the two weeks. This real-world evidence validated the barge-in-agency hypothesis from the WoZ and convinced the engineering team to prioritize building it for the next iteration of the product.

The feedback from drivers who used the MVP in front of riders was mixed. Some were neutral or even excited to use it in a social context. Others described it as “awkward” or mentioned being naturally shy which led them to disuse the feature in that situation. This individual variability reflects drivers’ judgement of agency which was discussed in the literature review - that is agency is not simply a physiological phenomenon but one also influenced by a person’s context information, background beliefs and social norms. Because these traits vary by person, some drivers did not actually experience reduced agency from using the VUI in front of people while others did. Given that drivers still had agency in using the touch modality to accept or decline trip requests, the team decided the engineering effort to disable the feature in a social context was not necessary for the wider release of the MVP. Nevertheless, building the ability to turn the VUI on and off depending on the situation remains a goal for the team in the long-term.

Analogous to the impact of individual variability on perceived agency, it was learned that driver perception of the VUI’s personality could also impact agency, a factor not identified during the WoZ. This ascription of personality to an artificial agent is a widely recognized phenomenon known as anthropomorphism. Given VUIs transmit information in natural language, demonstrate contingent behavior and play a social role through autonomous assistance, they are especially likely to elicit anthropomorphic responses from humans (Reeves & Nass 1996). In fact, multiple studies find increased customer satisfaction in conversational interfaces when they are more human-like (Araujo 2018; Waytz 2014,). To mitigate the risk of users perceiving a VUI’s personality in unintended ways, it’s recommended designers invest time in formulating and testing a persona for it. In the case of the Uber MVP VUI, the team aimed for a neutral tone but made no significant effort to intentionally formulate its persona due to lack of resources. Nevertheless, the alpha research revealed drivers negatively perceived the MVP when declining trip requests. Specifically, they reacted to it saying “*okay, we will let this one pass*” after declining a trip request. Drivers described the phrase variously from “sarcastic” and “snarky” to “stern” and “authoritative” alluding to a perceived imbalance in the power dynamic between them and the VUI. Despite the reality of having full control, drivers faced a reduction in their proxy and collective agency by feeling less able to influence and work with the VUI. As a result of this finding, the team changed the utterance to simply say “*trip declined*” and more importantly onboarded a VUI designer to initiate a separate line of research to define and validate an intentional persona.

Two additional studies conducted in India and Australia months after the alpha research confirmed the previous findings in different cultural contexts and validated the value of barge-in which was ready on iOS devices. In India, reduced agency remained an issue due to the lack of barge-in functionality on Android devices which all participants used. In Australia, all participating drivers used the iOS version with barge-in and reported increased agency and satisfaction. One driver described his experience using barge-in saying “*A good thing I noticed is the minute it came up, you could say yes, you didn't have to wait.*” As far as using the MVP in front of riders, drivers in Australia had mixed feelings similar to those in the US. In the case of India, the majority of drivers reported not being shy about using it in front of riders. While the reason for this disparity is unclear, it raises a larger question of how different cultural norms might influence agency in human-AI interactions - a topic ripe for exploration. Finally, since changes to the VUI's decline trip response had been made before the studies in India and Australia, no drivers perceived its persona negatively.

SELLING STAKEHOLDERS ON AGENCY

The WoZ and alpha research established agency as an important factor for the VUI's success and surfaced the need for specific features to foster it, especially barge-in. To convince product and engineering stakeholders to prioritize it, three strategies were employed. First, video of five drivers interrupting during the WoZ multiple times was stitched together and shown to the product and engineering teams. Additional video captured during the US alpha research reinforced barge-in's potential value to the user experience and helped convince the team to commit to building it. The use of video as both a storytelling and evidence device proved more powerful than the researcher reading driver quotes or telling what happened and is recommended for practitioners trying to convince teams of the importance of fostering agency in their products. Best practices learned include keeping videos brief to hold stakeholder attention, using evidence from multiple users and ensuring the videos are shareable so teams can rewatch and share them to a broader audience beyond formal presentations.

The second convincing strategy connected agency to product usage - a business outcome all product stakeholders care about. Figure 5 shows the actual slide presented to the product and engineering teams to argue for barge-in framing it as “supercharging” the feature with multiple discrete benefits like reduced waiting anxiety and a lowered learning curve. Furthermore, agency was referred to as control, a more easily understood concept. Framing barge-in in this way connected the VUI's success to agency and helped spur the team to commit to building it. Practitioners are encouraged to similarly tie agency to business outcomes like product usage to more easily convince stakeholders of its importance.

The third strategy was to directly involve software engineers and other product team members in the research. For example, during the alpha phase, three engineers participated in semi-structured interviews with drivers hearing first-hand how reduced agency affected their experience with the product. In one instance, a driver described their experience to an engineer saying “*I wasn't really sure when it was working and started to panic.*” Though engineers do not typically participate in user research at Uber given their imperative to focus on coding, including them in the research proved especially effective at convincing them to prioritize agency-supporting features like barge-in. One lightweight approach to encourage engineer participation was inviting them to observe interviews through remote video-conference. To

do this, the researcher placed optional invites on engineers' calendars and when a session started, posted the video-conference link to the team's internal messaging channel as a reminder. In one instance, an engineer, who had previously participated in a session, strongly urged his colleagues in the team messaging channel to view them. The researcher also invited the engineering team to a small subset of the scheduled interviews based on their potential to generate interesting insights which the researcher tried to predict from participant screener responses.

Barge-in will supercharge feature

Insight
"I'm anxious...waiting for prompt so not to miss trip opportunity. Once I get a trip offer, I start saying Accept, Accept, Accept until she responds." - SF Partner

Why Barge-in?

- It's analogous to the learned behavior of immediately tapping a dispatch
- It reduces waiting anxiety
- It gives control to the user
- It lowers the learning curve
- It allows for longer, more helpful prompts

A photograph of a man in a dark suit and tie, smiling and looking towards the right. He is standing next to a dark-colored car. The background shows a city street with buildings and a fountain.

Figure 5. Slide presented to product and engineering teams to argue for barge-in.

DISCUSSION

This case described a novel implementation of the WoZ method which uncovered agency as a critical mediator of driver interaction with a proposed VUI for Uber's driver application, presented three factors that impacted agency and provided learnings on how to convince stakeholders of the importance of agency to the user experience. For the WoZ, the in-car setting and lack of a truly functioning prototype required the researcher to act as both wizard and facilitator and slightly deceive the participant about the nature of the setup. While deception is inherently necessary to run a WoZ, researchers should evaluate the ethical implications regardless of the low risk of not telling participants they're in a simulated experience in most cases. The WoZ is also gaining respect as a viable ethnographic tool and proved invaluable in identifying the importance of agency in the Uber VUI project. This case study along with the use of the WoZ in other recent ethnographic work like Osz and Stayton's studies on autonomous vehicles (Osz 2018; Stayton 2017) show it should be seriously considered by researchers exploring human-AI products. In the case of the Uber research team, this was the first use of a WoZ to evaluate futuristic technology. Barriers to earlier adoption were primarily technical with limited engineering resources available to create high-fidelity wizard interfaces for research only. The logistics of running a WoZ in a

moving vehicle further raised the barrier to entry. Nevertheless, the use of the WoZ in this case piqued the interest of some researchers on the team with one subsequently conducting an in-car WoZ focused on maps and navigation. Researchers in industry are encouraged to think of creative ways to implement WoZs if facing similar limitations.

Three factors were found to impact driver agency during the WoZ and alpha research: conversational agency, usage in social contexts and perception of the VUI persona. To exercise conversational agency, drivers interrupted the VUI and reduced their usage of the accept-decline MVP when barge-in functionality was not supported. Interestingly, the need for conversational agency was greatest during VUI-initiated interactions. While the reason behind this was not explored in this research, it likely relates to a lack of shared agency established between driver and VUI. As mentioned in the literature review, collective or shared agency is essential for groups to complete objectives together successfully. In this case, drivers did not feel that the VUI was a true partner yet. Fostering such shared agency will be critical to successful human-AI collaboration and practitioners in industry are urged to think deeply about how to do so in their products. For the Uber VUI, the team is considering future features like voice personalization. For drivers uncomfortable with speaking to the VUI in a social context, agency was reduced leading to disuse of the accept-decline MVP in front of passengers. Thus, companies are strongly advised to consider how the different contexts in which their AI products will be used might impact agency and subsequently user behavior. Finally, negative perception of the VUI persona when declining trip requests caused drivers to feel like they did not have influence over the VUI (proxy agency) or that it was a collaborative partner for them (collective agency). Organizations should thus consider how users will perceive their conversational interface products and strive to intentionally design and validate personas for them.

In general, researchers and designers of AI technology should account for agency early in the product development process. While it can seem abstract and inconsequential, the ramifications of ignoring it can directly impact the user experience and eventual business outcomes. Paying attention to agency early also gives researchers ample time to convince product and engineering teams of its importance. To convince these stakeholders, this case suggests researchers use video, connect agency to business outcomes and involve engineering and product directly in research. Ultimately, this case study underscores the need to better understand how to foster agency in emerging AI technology. Drivers showed they would reduce their usage of the Uber VUI when they felt diminished agency. And while some insight into how to mitigate this in conversational interfaces specifically was provided, whole new interaction paradigms and strategies will be needed for the wide range of advanced AI technology on the horizon.

CONCLUSION

As AI technology becomes smarter and more capable, individual agency will yield to a more collective one shared between human and machine. This reconfiguration begs many questions. Will people willingly accept less control? How might people and AI work well together? How might designers foster agency sharing? This case begins to answer some of these questions. It describes the use of the WoZ method as a valid ethnographic means to evaluate agency in proposed technology pre-development, provides insight on factors that impact agency in VUI interactions and details three strategies for communicating the

importance of agency to a cross-functional product team in a major tech organization. Most importantly however, it demonstrates that agency matters. Drivers showed they would not accede to the role of subordinate and disused the VUI when they felt less agency in their interactions with it. Designers of AI technology should take note and strive to balance the agency embedded in the logic and algorithms of their creations with the need for users to have their own agency. Erika Stayton sums up this tension perfectly saying there's a "*strange polysemy at the heart of autonomy: one may be freed from certain tasks but also further embedded in sociotechnical systems that are beyond individual control*" (Stayton 2017). Researchers and ethnographers in industry are uniquely placed to bring a humanistic perspective to the development of these sociotechnical systems and inspire their organizations to foster agency in their newfound technological wonders. Doing so is business critical.

Jake Silva is a User Researcher at Uber. His research focuses on conversational interfaces, customer care products and log-based quantitative methods. He received his Master of Science in Information from the University of Michigan School of Information and Bachelors from American University in Washington, DC. Email: silva@uber.com.

NOTES

Maria Cury, Elena OCurry and George Zhang are thanked for their invaluable feedback on improving this case for the EPIC audience. Daier Yuan was also a critical partner in assisting with planning and implementing the Wizard-Of-Oz research in New York City and Chicago.

REFERENCES CITED

Appel, Jana, Astrid von der Pütten, Nicole C. Krämer, and Jonathan Gratch. "Does humanity matter? Analyzing the importance of social cues and perceived agency of a computer system for the emergence of social reactions during human-computer interaction." *Advances in Human-Computer Interaction* 2012 (2012): 13.

Applin, S., & Fischer, M. (2015). Cooperation between humans and robots: applied agency in autonomous processes!. In 10th ACM/IEEE International Conference on Human/Robot Interaction, The Emerging Policy and Ethics of Human Robot Interaction workshop, Portland, OR.

Astrid, M., Nicole C. Krämer, Jonathan Gratch, and Sin-Hwa Kang. "It doesn't matter what you are!" Explaining social effects of agents and avatars." *Computers in Human Behavior* 26, no. 6 (2010): 1641-1650.

Araujo, Theo. "Living up to the chatbot hype: The influence of anthropomorphic design cues and communicative agency framing on conversational agent and company perceptions." *Computers in Human Behavior* 85 (2018): 183-189.

Bandura, Albert. "Exercise of human agency through collective efficacy." *Current directions in psychological science* 9, no. 3 (2000): 75-78.

Elish, M. C. "(Dis)Placed Workers: A Study in the Disruptive Potential of Robotics and AI." *We Robot* 2018

- Endsley, Mica R., and Esin O. Kiris. "The out-of-the-loop performance problem and level of control in automation." *Human factors* 37, no. 2 (1995): 381-394.
- Garza, Alejandro. 2019. "This AI Software Is 'Coaching' Customer Service Workers. Soon It Could Be Bossing You Around, Too." *Time Magazine* website, July 8. Accessed [October 25, 2019]. <https://time.com/5610094/cogito-ai-artificial-intelligence/>
- Gibson, David R. "Seizing the moment: The problem of conversational agency." *Sociological Theory* 18, no. 3 (2000): 368-382.
- Haslbeck, Andreas, and Hans-Juergen Hoermann. "Flying the needles: flight deck automation erodes fine-motor flying skills among airline pilots." *Human factors* 58, no. 4 (2016): 533-545.
- Heer, Jeffrey. "Agency plus automation: Designing artificial intelligence into interactive systems." *Proceedings of the National Academy of Sciences* 116, no. 6 (2019): 1844-1850.
- Holstein, Kenneth. "Towards teacher-ai hybrid systems." In *Companion Proceedings of the Eighth International Conference on Learning Analytics & Knowledge*. 2018.
- Johnson, Gregg. 2017. "Your Customers Still Want to Talk to a Human Being". *Harvard Business Review* website, July 26. Accessed [October 16, 2019]. <https://hbr.org/2017/07/your-customers-still-want-to-talk-to-a-human-being>
- Manyika, James, Michael Chui, Mehdi Miremadi, Jacques Bughin, Katy George, Paul Willmott, and Martin Dewhurst. 2017. "Harnessing automation for a future that works" *McKinsey Global Institute* website. Access [October 14, 2019]. <https://www.mckinsey.com/featured-insights/digital-disruption/harnessing-automation-for-a-future-that-works>.
- Milanesi, C. 2016 "Voice Assistant Anyone? Yes please, but not in public!. Creative strategies", 3 June, accessed 25 June 2017. <https://creativestrategies.com/voice-assistant-anyone-yes-please-but-not-in-public>.
- Moore, James W. "What is the sense of agency and why does it matter?." *Frontiers in psychology* 7 (2016): 1272.
- Moorthy, Aarthi Easwara, and Kim-Phuong L. Vu. "Voice activated personal assistant: Acceptability of use in the public space." In *International Conference on Human Interface and the Management of Information*, pp. 324-334. Springer, Cham, 2014.
- Osz, Katalin, Annie Rydström, Vaike Fors, Sarah Pink, and Robert Broström. "Building Collaborative Test Practices: Design Ethnography and WOz in Autonomous Driving Research." *IxD&A: Interaction Design and Architecture (s)* 37 (2018): 13.
- Roose, Kevin. 2019 "A Machine May Not Take Your Job, but One Could Become Your Boss" *New York Times* website, June 23. Accessed [October 16, 2019]. <https://www.nytimes.com/2019/06/23/technology/artificial-intelligence-ai-workplace.html>
- Sahaï, Aïsha, Elisabeth Pacherie, Ouriel Grynszpan, and Bruno Berberian. "Co-representation of human-generated actions vs. machine-generated actions: Impact on our sense of we-agency?." In *2017 26th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*, pp. 341-345. IEEE, 2017.

Shneiderman, Ben, Catherine Plaisant, Maxine Cohen, Steven Jacobs, Niklas Elmqvist, and Nicholas Diakopoulos. *Designing the user interface: strategies for effective human-computer interaction*. Pearson, 2016.

Simon, Matt. 2019. "Robots Alone Can't Solve Amazon's Labor Woes" WIRED website, July 15. Accessed [October 14, 2019]. <https://www.wired.com/story/robots-alone-cant-solve-amazons-labor-woes/>.

Sloane, Steve. 2019. "The shift to collaborative robots means the rise of robotics as a service" TechCrunch website, March 3. Accessed [October 14, 2019]. <https://techcrunch.com/2019/03/03/the-shift-to-collaborative-robots-means-the-rise-of-robotics-as-a-service>.

Stayton, Erik, Melissa Cefkin, and Jingyi Zhang. "Autonomous Individuals in Autonomous Vehicles: The Multiple Autonomies of Self-Driving Cars." In *Ethnographic Praxis in Industry Conference Proceedings*, vol. 2017, no. 1, pp. 92-110. 2017.

Sun, Y., & Sundar, S. S. (2016, March). Psychological importance of human agency how self-assembly affects user experience of robots. In *2016 11th ACM/IEEE International Conference on Human-Robot Interaction (HRI)* (pp. 189-196). IEEE.

Synofzik, Matthis, Gottfried Vosgerau, and Martin Voss. "The experience of agency: an interplay between prediction and postdiction." *Frontiers in psychology* 4 (2013): 127.

Waytz, Adam, Joy Heafner, and Nicholas Epley. "The mind in the machine: Anthropomorphism increases trust in an autonomous vehicle." *Journal of Experimental Social Psychology* 52 (2014): 113-117.