Privacy, Power, and Invisible Labor on Amazon Mechanical Turk

Shruti Sannon
Cornell University
Ithaca, New York
ss3464@cornell.edu

Dan Cosley
Cornell University
Ithaca, New York
drc44@cornell.edu

ABSTRACT

Tasks on crowdsourcing platforms such as Amazon Mechanical Turk often request workers' personal information, raising privacy risks that may be exacerbated by requester-worker power dynamics. We interviewed 14 workers to understand how they navigate these risks. We found that Turkers' decisions to provide personal information during tasks were based on evaluations of the pay rate, the requester, the purpose, and the perceived sensitivity of the request. Participants also engaged in multiple privacy-protective behaviors, such as abandoning tasks or providing inaccurate data, though there were costs associated with these behaviors, such as wasted time and risk of rejection. Finally, their privacy concerns and practices evolved as they learned about both the platform and worker-designed tools and forums. These findings deepen our understanding of both privacy decision-making and invisible labor in paid crowdsourcing, and emphasize a general need to understand how privacy stances change over time.

CCS CONCEPTS

• **Information systems → Crowdsourcing**; • **Security and privacy → Human and societal aspects of security and privacy**;

KEYWORDS

Privacy, invisible labor, crowdsourcing, MTurk

ACM Reference Format:
Shruti Sannon and Dan Cosley. 2019. Privacy, Power, and Invisible Labor on Amazon Mechanical Turk. In *CHI Conference on Human Factors in Computing Systems Proceedings (CHI 2019), May 4–9, 2019, Glasgow, Scotland UK*. ACM, New York, NY, USA, 12 pages. https://doi.org/10.1145/3290605.3300512

CHI 2019, May 4–9, 2019, Glasgow, Scotland UK

ACM ISBN 978-1-4503-5970-2/19/05…$15.00
https://doi.org/10.1145/3290605.3300512

1 INTRODUCTION

Digitally-mediated work is becoming increasingly common, with almost one in every ten Americans earning money via an online work platform [31]. Many platforms require workers to disclose personal information, both up front (as with worker profiles in Upwork) and while working (as with Uber drivers' locations). On Amazon Mechanical Turk (MTurk), a popular crowdsourcing site, requesters can post tasks ("Human Intelligence Tasks", or "HITs") that ask workers ("Turkers") to provide personal demographics, answers to deeply personal questions, or even video of themselves.

Although issues of disclosing personal information are not unique to MTurk or to online work, they may be magnified in digital labor due to stark information and power asymmetries. Workers often have limited information about requesters or the content and purpose of a given HIT, and have little recourse if requesters reject their work [23]. In contrast, requesters can aggregate Turkers' personal information by asking for different types of data across multiple HITs [14]. These factors may make it hard for workers to accurately assess the privacy risks of completing any given task. Digital work is also often precarious and does not offer traditional labor protections, increasing risk to workers while depressing their wages [30]. This economic power imbalance may also compromise workers' privacy; although Turkers have higher privacy concerns than average [15], they may discount legitimate privacy concerns to earn needed income or to avoid consequences such as being blacklisted [29].

These considerations make MTurk, and digital labor more generally, a compelling context for studying both privacy concerns and privacy-protective behaviors (PPBs). Turkers have reported many privacy concerns and violations around data collection and profiling, unauthorized use of data, invasive stalking and spamming, and deceptive practices such as phishing and scams [36]. How Turkers navigate these risks and make decisions about disclosing their personal data during the course of their work, however, is an open question that could both inform the design of crowd work ecosystems and deepen understanding of the unpaid work, or "invisible labor" [33], required to be an effective Turker.

Crowd workers' privacy and invisible labor are also topics close to home for CHI. Many HCI researchers design

crowd workflows or use MTurk in studies, and as we will see, Turkers tend to be more willing to disclose personal information to researchers than other requesters. MTurk is estimated to have 100–200k workers, with tens of thousands of new workers joining the platform every year [6]. Thus, we view reducing Turkers' privacy-related risks and labor as a practically and ethically important issue for the CHI community.

To understand how Turkers navigate privacy issues during their work, we conducted semi-structured interviews with 14 Turkers who had varying levels of experience and financial dependence on MTurk. We show that privacy considerations influence both what work gets done and how, as well as rationales for and costs of the PPBs Turkers use to protect themselves. We discuss how power, invisible labor, and time are all important lenses for understanding privacy and offer suggestions for how crowd work platforms, and the ecosystems of worker-created resources around them, can be designed to reduce both privacy risks and wasteful privacy-related invisible labor.

2 RELATED WORK

We begin by discussing how Turkers select tasks, highlighting the invisible labor associated with this process. We then examine the privacy risks and concerns experienced by workers, and the gaps in knowledge our study intends to fill around practices, power, and privacy in choosing and doing HITs.

Choosing and Working on HITs: A Walkthrough

A key decision for Turkers is which HITs to work on, and though Turkers work for many reasons, including fun, the primary motivation is money [16]. Thus, pay rate is a key concern, along with the novelty, speed, and repeatability of the HIT [18]. Workers sort through HITs to find new, high-paying HITs [4], often using Turker-written browser scripts to help find and capture them.

Perceptions of requesters also affect Turkers' decisions about HITs. There are many different types of requesters, including academic institutions and private companies; Turkers develop opinions about both individual requesters and types of requesters as they work. Savvy Turkers vet individual requesters by both writing and checking reviews left by other Turkers on sites such as Turkopticon [13]. Issues such as clarity of communication and task design [23] and fairness around rejection [25] are important considerations in evaluating requesters. Privacy issues are less often talked about, though exploratory work suggests that academic requesters are seen as more trustworthy [29].

The type of work involved in a HIT can also affect Turkers' decision-making. Surveys are the most popular HITs among U.S. workers [7], but HITs can involve other types of work, including searching the Internet for information, verifying information, interpreting or categorizing data, creating new content, and clicking links to access content [7, 9]. Some HIT types may pose more privacy issues than others. For example, surveys can collect sensitive personal information, while content creation tasks can ask workers to take pictures or recordings of themselves.

Turkers also organize communities, such as MTurk-related Reddit forums, TurkerNation, and MTurkForum, to communicate about requesters, HITs, scripts, and Turking. About 60% of U.S. workers report using such forums [38], which play an important role in knowledge sharing and reducing task completion times [37]. Workers also share information about good requesters and high-paying tasks with their networks, giving committed, connected Turkers an edge [38].

Overall, Turkers engage in a large amount of invisible labor—that is, work to complete HITs that is unpaid [23]—around searching for HITs, working on HITs that are rejected, and beginning to work on HITs but then choosing to abandon them ("returning" HITs) [12]. Returning HITs is common: through an analysis of 2,676 workers and 3.8 million HITs, Hara et al. found that 12.8% of all HITs were returned [12]—wasted effort that significantly reduced average hourly wages. We suspect that many of these HITs are returned due to privacy concerns about personal information.

Risks and Privacy Issues on MTurk

Turkers are supposed to be anonymous, and MTurk's Acceptable Use Policy prohibits collecting personally identifiable information (PII) such as email addresses [35][1]. Still, many tasks require Turkers to provide personal information without a guarantee of confidentiality [8], and Turkers have reported quite invasive requests, such as for photographs of their health insurance cards [36]. They also express concerns about how their information is collected and used, as well as malicious requesters who may spam or scam them [36].

Further, individual requests for small amounts of personal information can lead to privacy risks through aggregation. Kandappu et al. launched three seemingly unrelated HITs to collect individually innocuous PII from workers, using MTurk IDs to connect data from the individual HITs to profile a Turker's birthday, gender, and zip code [14]. Easy access to this information is problematic, as research using U.S. census data indicates that these three data points alone can be used to identify 87% of the U.S. population [34]. MTurk IDs are also linked to workers' Amazon accounts, meaning workers can often be identified through a regular Internet search [20].

Finally, although Turkers are more privacy-conscious than the general population [15], they may not always know the

[1]Turkers generally refer to HITs that request PII as Terms of Service (TOS) violations, as will we in the remainder of the paper.

risks posed by the site. On balance, they care about remaining anonymous, and most believe (incorrectly) that a requester cannot find out their full name [20]. They are also generally unaware that requesters can profile them across tasks, though most would not knowingly complete HITs that aim to profile them [14]. About a third of U.S. Turkers surveyed did not report any privacy concerns or negative privacy-related experiences; whether these Turkers are generally unconcerned about privacy, unaware of the risks, or prioritize earning money over privacy is an open question [36].

Putting It All Together

On balance, research shows that Turkers put substantial effort into selecting and working on HITs and face real privacy risks while doing them. However, research has not examined how privacy risks influence Turkers' decisions around what HITs to accept and to complete, nor how they respond to requests for personal information. Research on other online contexts, such as social media or online shopping, finds that people engage in a privacy calculus, weighing the benefits of providing their information against the costs to their privacy [5]. We suspect that this privacy calculus takes on a new dimension on MTurk because decisions about which HITs to complete have real consequences for Turkers' income. Similarly, people in other contexts also engage in a range of PPBs to address privacy concerns [32]; however, it remains to be seen when and to what extent Turkers engage in PPBs, and how the uneven power dynamics and financial dependence on MTurk influence the degree to which they are able to protect themselves on the site.

3 METHODS

To understand how privacy affects Turkers' decision-making and practices, we conducted semi-structured interviews with 14 Turkers about how they experience working on the site, focusing on issues around privacy and requests for personal information.

Recruitment and Procedure

We recruited U.S.-based participants at least 18 years of age. We originally posted the study as a HIT on MTurk itself, but got no responses, we think because high-paying requests from new requesters are viewed with some suspicion. So, we advertised the study on two popular MTurk forums: /r/mturk, a Reddit forum for Turkers with over 40,000 subscribers as of January 2019, and TurkerHub (renamed TurkerView Forum in December 2018), another commonly used forum. We considered advertising on other sites as well, but the responses we got were diverse enough that posting on additional forums felt unnecessary.

On both sites, we posted a new thread advertising the HIT, respecting community rules about how requesters can post recruitment messages. The HIT was advertised as "an interview about Turkers' experiences" to avoid only recruiting Turkers with strong feelings about privacy. Interested Turkers contacted us via private messages on the forums, and after optionally viewing the IRB-approved consent form on our institution's official website to verify our identity as academic researchers, chose an interview time and were granted access to the study through a qualification (i.e., an exclusive HIT) on MTurk. After accepting the HIT, participants completed the consent form, then received a link to a secure, anonymous chatroom to chat with the first author. Using a text-only, anonymous chatroom allowed us to avoid collecting PII such as email, Skype IDs, or audio. Interviews were conducted between May and August of 2018.

The first author conducted all interviews one-on-one based on a semi-structured interview guide. We first asked participants about their overall MTurk experience, including why they started Turking, types of HITs they enjoyed and disliked, and how they decided which HITs to do and requesters to work for. We then asked about the kinds of personal information HITs demand and their thoughts about providing such information. We used the term "personal information" rather than "privacy" at first to avoid activating privacy framings that might shape participants' responses. We then asked more explicitly about participants' perceived privacy risks, how they make decisions about revealing information during HITs, and why they do or don't engage in PPBs. Finally, we asked about the role Amazon and other Turkers have played in their MTurk experience and how that experience, including privacy perceptions, changed over time.

The guide was developed based on our research questions and was informed by participant-observation, a valuable practice for crowd work research [1]. The first author worked on 637 HITs over the course of a month to develop a sense of the decisions involved from a Turker's perspective, recording field notes [21] to track emerging issues. These observations helped us probe further during the interviews (e.g., about participants' attitudes about specific types of tasks, such as webcam HITs). Talking about the first author's own Turking also helped build rapport with participants and mitigate requester/worker power dynamics. Interviews lasted about 60 minutes, after which participants completed a brief demographic survey and were compensated $15. No personally identifiable information, including MTurk IDs, was stored.

Analysis

The first author wrote memos after each interview to capture main points and identify new questions to add to the interview guide [3]. For example, the first few interviews suggested time would be a useful lens for understanding Turkers' privacy experiences, leading us to add questions

about how participants' privacy concerns and MTurk experiences had evolved.

We used the constant comparative method to see whether emerging themes differed across groups of Turkers (e.g., full-time versus part-time, or new versus experienced); this allows researchers to identify patterns in the data and ensure theoretical saturation [10]. The co-authors discussed the emerging themes on an ongoing basis. We reached theoretical saturation at ten interviews, after which no new concepts emerged. We conducted subsequent interviews to confirm and deepen the initial analyses.

We conducted a qualitative interpretive analysis of the transcripts [21]. All transcripts were read by both coauthors. Then, the first author assigned open codes to a random subset of 5 transcripts and both authors met to discuss the codes and connections between them and develop focused categories and themes. For example, codes related to PPBs were organized into two main groups: type of PPB and rationales for engaging in them, and reasons not to engage in PPBs. The first author then assigned focused codes to the transcripts in NVivo. As with the interviews, both co-authors regularly discussed the development and analysis of codes, categories, and themes.

Participant Characteristics

As shown in Table 1, participants ranged from 22 to 65 years old, with an average of 35 years. Our sample was gender-balanced, with six participants identifying as male and eight as female. Twelve identified as Caucasian and two as being multiracial. Education levels included a postgraduate degree (1), some graduate school (2), a bachelor's degree (1), some college (8), and a high school diploma (2).

We sought participants with varying levels of economic dependence on MTurk to probe how economic necessity drove their choices. Four relied on MTurk for their primary income (e.g., MTurk income made up their rent), while ten used MTurk to supplement primary income from another source (e.g., MTurk income went towards necessities, such as groceries, or pocket money, such as movie tickets). Of these ten, six were employed full-time, two part-time, one was retired, and one was a full-time student. Most (11) reported a yearly household income between $20,000 and $50,000. Most (9) lived in urban areas, with three in suburban and two in rural areas. They had between five months and seven years of experience on MTurk, with a median of 17,280 completed HITs across a variety of work types, including transcriptions, surveys, and audio/video recordings. Collectively, participants had completed almost one million HITs.

Table 1: Participant Demographics

ID	Gender	Age	MTurk Income*	Experience
P1	F	35	Supplementary	9 months
P2	M	36	Supplementary	7 years
P3	M	24	Supplementary	1.5 years
P4	M	33	Supplementary	6 years
P5	F	51	Primary	1 years
P6	M	65	Supplementary	9 months
P7	F	34	Supplementary	5 months
P8	F	31	Supplementary	7 months
P9	F	28	Supplementary	5 years
P10	F	39	Primary	2.5 years
P11	M	36	Primary	3 years
P12	F	23	Supplementary	5 years
P13	M	22	Primary	1 year
P14	F	37	Supplementary	5 months

*Whether MTurk provides primary or supplementary income

Limitations

Our study has several limitations that we outline to help contextualize and interpret our findings. First, we recruited from two popular forums for MTurk workers, excluding workers who do not know about or do not use these forums. We don't see this as a showstopping limitation: prior work suggests that 60% of U.S. workers communicate in forums [38], our participants described their MTurk experience before they started using the forums, and our sample broadly mirrors other studies of MTurk demographics in terms of gender, age, and income [6]. But it does mean that we oversampled from more experienced workers who sought out information and advice on how to be effective, and we miss Turkers who have left the platform, some of whom may have left due to privacy-related concerns.

We also restricted our sample to U.S. Turkers, about 75% of all workers [6]. Thus, our findings cannot speak to other cultures, such as the 16% of Turkers who live in India [6]. We know there are cultural differences in privacy concerns between Indian and U.S. Turkers [36] and that Indian Turkers face structural barriers including higher administrative overheads and being unwelcome in U.S. MTurk forums [11]. Future work should explore possible cultural differences.

Finally, our sample is limited to people who were comfortable with being interviewed, and thus may miss more privacy-conscious Turkers. Conducting the interviews via anonymous chat may have mitigated this, as several participants said they would not have participated in an audio or video interview because these are more identifying media. While chat took more time than a verbal interview, the use of informal language helped reduce the distance between the interviewer and participants, and the delay in interaction helped us formulate careful follow-up questions. Moreover, while people may write less than they would speak on a subject, participants often responded in full paragraphs and

appreciated the chance to talk about their experiences: one participant referred to the interview as "MTurk therapy."

4 FINDINGS

We organize our findings into three main themes: 1) factors that influence how Turkers judge the acceptability of requests for personal information, 2) the reasons why Turkers engage in or eschew a range of privacy-protective behaviors, and 3) how gaining experience influences Turkers' privacy-related processes. Findings we report below are based on codes and themes that occurred frequently in the dataset.

Judging Requests for Information

Several of the same factors that affect Turkers' decisions about HITs in general—including pay, attributes of requesters, and the nature of the HIT—influenced how Turkers perceived requests for personal information and how they navigated privacy-related issues on the site.

Economic Bottom Line. Judging the acceptability of requests for information was a balancing act that often was based on pay: "Make money while trying to maintain some privacy. It's a tightrope" (P4). A certain amount of money was worth complying with requests for information, as P11 explained: "There's an amount of pay that makes it worth it to share, an amount that makes it worth it to fudge the name and date just a little, and an amount that isn't worth bothering with at all." Money also justified taking on some risk: "there were one or two HITs that asked for my full name that I did anyway. [they] paid pretty well" (P7).

Money could sometimes trump all other factors, in line with prior findings that Turkers' primary motivation is to earn needed income [16]. For example, P3 described a suspicious but well-paying HIT: "I had a requestor ask me to pour ketchup on my chest while lying down in my bathroom. Can't imagine what that was for, and I didn't ask. [...] it's probably on some creep site somewhere." When asked why he completed the HIT, he explained: "I thought about it in terms of true economics: cost-benefit analysis. It takes me very low effort and only a couple minutes to prepare and film it. Second, it was already in the bathroom so clean up was quick. I finished it in about 15 minutes total so that was around 12–15 bucks an hour pay rate." However, other participants were less willing to put a dollar amount on their privacy: "if you feel uncomfortable giving out information, dont give it out. it's just a survey and not worth the 50 cents or $1 they're giving you" (P8).

Attributes of the Requester. Almost all participants considered requester characteristics when evaluating privacy risk. Requesters from academic institutions were widely trusted, since they were perceived to have been vetted by an IRB: "I find some 'safety' with the University based HITs as they usually have more rules to follow from the IRBs" (P1).

Academic requesters were also perceived as having a legitimate purpose for data collection, which set them apart from private companies. Participants also liked that academic requesters were answerable to an IRB, providing some recourse: "I would not even consider answering highly personal questions for a non-university affiliated requester. Have you seen the foot fetish guy on MTurk? People are weird LOL. At least with university studies I know there's some accountability via their IRB, and *hopefully* a legitimate scientific purpose to asking the questions they ask" (P13).

However, being an academic requester wasn't enough; HITs still had to pay enough to justify the privacy costs, as P10 noted about a decision to provide her mailing address to Harvard. Further, Turkers were aware that unscrupulous requesters could pose, phishing-style, as academics: "The thing is...how do you know it's a legit university requester? I could get a Cornell agreement and slap it on my HIT, then scam people. MTurk has no safeguards for workers" (P4).

Requesters could also build trustworthy reputations on MTurk forums: "The great requesters are well known on the forums... these are the ones I usually work on. There are many established requesters we all know about on the forums" (P2). Participants also used Turkopticon (often abbreviated in text as 'TO') and TurkerView to evaluate requesters, and Turkopticon reviews allow Turkers to flag a HIT as a potential Terms of Service (TOS) violation: "On TO, users can designate if a HIT violates the TOS, and I tend to steer clear of those if I think my personal files/accounts could be accessed somehow" (P9).

That said, some Turkers were skeptical of Turkopticon ratings. For instance, P1 did not trust them because they are written by other Turkers who are often in direct competition: "For the most part good requestors show up in the Green on TO or thumbs up on TV. The great requestors are often in the Red or not ranked at all. Many Turkers won't give away their great ones for fear the work will get scooped up from them" (P1). Participants who shared these views had started to disregard Turkopticon ratings, relying on their intuition instead: "I try to be honest and always put positive reviews on requesters I enjoy but it's hard to rely on TO anymore. I've been on here for a long time to be able to tell which requesters would be scamming" (P12). A few of these participants had migrated to TurkerView: "TV has more moderation, and 'bad data' will be removed" (P13). This work of finding, evaluating, and writing reviews constituted additional invisible labor that was necessary to accurately assess requesters.

Purpose or Creepiness of the HIT. Participants also evaluated individual HITs. Avoiding scams was seen as not too hard: "There are always a few HITS that try to get you to sign

up for referral link services, 'Sign up and earn $10 credit'. You can tell easily these are phishing for emails and referral. There are many Cryptocurrency sites people want us to sign up for. A seasoned Turker will spot them easily" (P2). HITs could also become suspicious halfway through: "I would finish one that looks normal, then come to the demographics page that wants my address, full name and phone number to send me my 'results' "(P12). Finally, some HITs felt "creepy" even if they weren't outright scams: "when I first started there was this HIT asking for pictures of eyes. It paid decently for the work, I guess, and I thought about it but then I was like nope... that's creepy to me for some reason" (P14).

The perceived purpose of a request also affected how participants evaluated it. Participants were used to providing basic information such as age, gender, and zip code; these requests were seen as acceptable since, as P3 explained, "it's just for classification purposes." Most were also comfortable providing other types of personal information if the request seemed for a legitimate purpose: "I don't mind any of the typical demographic questions, or even if the personal question seems to have a clear link to the survey" (P1).

Some participants also considered risks not just to their own data but to third parties whose data was part of a HIT: "There was one when I first started out that was some kind of... I don't know, workers comp investigation company or something? They gave you the full name of a person and then some kind of external link (say, to a Twitter account) and had you check the link for evidence that the person had been active during a certain time period. It was really creepy" (P14). Lasecki et al. found that Turkers can be used in malicious tasks, such as extracting a credit card number from a photo, but that relatively few are willing to complete such a HIT when it seems malicious than when it appears innocent [19]. Similarly, although our participants avoided HITs that were clearly malicious, they did complete HITs that seemed more innocuous, even if they weren't sure whether the third parties had consented to their data being used, such as HITs involving rating people's dating profiles: "If I found out [the requester was] doing something uncool with the pictures, I'd definitely feel culpable, but I think the onus ultimately falls on the requester" (P9).

Risks of Compliance and Sensitivity of Data. Participants also assessed the degree of risk involved in complying with any given request, in terms of whether they could be identified, the sensitivity of the information requested, and the consequences of the information being misused or leaked. Requests for demographic information were common and not seen as particularly identifying; all participants were comfortable complying with these requests: "it's a bit mind-numbing providing the same demographics in HIT after HIT after HIT [...] They never ask for specifics that would be able to identify me, or at least that's how I think of it" (P6).

Some surveys asked Turkers to share personal opinions and life events, which participants generally didn't mind doing anonymously: "I have done studies and been totally honest on everything from depression, medications, sex life. as long as there is no name attached, I am okay with it" (P2). Some were willing to provide such sensitive information in text, but drew the line at audio or video, since these were identifying and thus additionally sensitive in a way that survey responses were not: "The only time I will stop is when they ask for Webcam access or sometimes access to microphone. I don't do the voice HITS either. I have never used my voice. I think that is the best way I can protect my unique personality from getting online" (P2). Participants also worried that video or audio could be used in unanticipated or (as with the ketchup HIT) unsavory ways, although beliefs about how specific requesters would use the data sometimes mitigated these concerns: "I think when I give my voice to Amazon or Google it's going to put through their machinery and then get lose in an ocean of other voices" (P11).

Finally, some HITs asked for access to Turkers' social media accounts. All of our participants drew the line at this, since it was seen as identifying, too personal, and a violation of MTurk's TOS. Some participants considered social media accounts more identifying than video or audio: "I've done a couple of Webcam HITs where you just talk to the camera, I'm okay with that. I mean it's not running my footage through a database to find me...lol. Facebook and Twitter link to me personally, and my family" (P4).

The Strengths and Weaknesses of Various PPBs

Evaluations of requests for personal information led Turkers to enact a range of PPBs around HITs, including avoiding risky HITs, returning HITs, telling privacy lies, withholding sensitive information, and reporting HITs. However, PPBs also posed costs, ranging from lost pay to fear of losing access to future work.

Avoiding Risky HITs. When possible, participants avoided accepting HITs they saw as asking for information that was risky or too personal. For example, requests to upload webcam footage often indicate this in the HIT title, allowing Turkers to avoid them if desired: "I avoid HITS that want to access my webcam or want an image of me" (P2). However, many HITs do not list the types of data required to complete them: "[it] happens far too frequently when it's not disclosed beforehand. I've gotten hits where it seems simple and a regular type of study then they ask for webcam access or my address" (P12). Thus, even if Turkers try to avoid certain types of requests, they may need to resort to other PPBs.

Returning HITs. The most commonly reported PPB was to return HITs when faced with concerning requests. This typically happened once participants had begun working on a HIT that later asked for personal information they were not willing to provide: "Researchers will often forget to mention they want you to take a selfie or to share a social media account and I'll definitely abandon most HITs right then and there" (P11). Sometimes, the issue was not about any one piece of information, but that the HIT required a lot of information overall: "if it's to where I think they're asking for too much, I get out of there quick" (P12).

However, returning HITs costs both money and effort: "I do return more than I should. It's frustrating and a loss, I will then block the requester so I don't waste my time again. It's one of the worst parts about working on MTurk" (P1). That said, participants generally saw returning a HIT as better than risking their privacy, with P13 describing the lost time in returning risky HITs as part and parcel of the MTurk experience: "usually you can tell pretty quickly something is a scam, so not much time will be wasted. However, sometimes you will have your time wasted but that's just the nature of MTurk. Just return the HIT and find something else to work on that is a valuable use of your time."

Privacy Lies. Some participants reported telling what Sannon et al. termed "privacy lies", providing inaccurate personal information to protect their privacy [28]. As in that study, these lies were often close to the truth, such as providing a birthdate that is only slightly inaccurate so as not to jeopardize researchers' data while protecting their identity and asserting their rights under MTurk's TOS: "if I'm asked for my actual birthdate I do fudge that. Technically I think that's a TOS violation to ask that" (P14). Unlike returning a HIT, privacy lies allowed participants to finish a HIT they had already spent time and effort on—and more wasted time justified more blatant lies: "if it's like ten minutes in and I've wasted my time I'll give a wrong phone number" (P4).

However, most participants were reluctant to tell privacy lies for both moral and practical reasons, also as found in Sannon et al. [28]. Many participants saw them as unethical, preferring to not provide any information: "I feel like if I am not comfortable to give that data [then] I shouldn't. I may do MTurk for money but I still try to be truthful" (P8). Participants also worried about harming academic researchers: "I never lie about demographic info—I did think about giving a fake name but I've never done that either. I guess because if I do the HIT and it's a survey I'm thinking about the poor grad student who just wants to do their research so I want to help them and give them good data" (P7).

Participants also feared that being caught could pose practical risks, including rejections or loss of access to work: "The best reason I have to share my real name is that some requesters come back again and again. If they ask for my name twice and 2 years ago I gave them a different fake name than I give them this time, they could know that and block me. If they block me, my account is at risk of being terminated by Amazon and then I'm out $35-40k/year" (P11). This highlights one way the power dynamics on MTurk intersect with privacy: workers are limited in terms of the PPBs they engage in, since requesters can cut off access to the market.

Selectively Sharing Sensitive Information. Another strategy participants used was to share only information they would not mind having discovered: "Even though I have revealed personal information in studies, I don't discuss anything that I would be horrified for people to know if it got out" (P14). Withholding sensitive information helped participants guard against data misuse or leakage: "I made sure that the demographics weren't too invasive. At most, they'd only know my gender and age if it did happen to be leaked" (P13). A related strategy to reduce harms was to only share information that people in offline networks already know: "anything I share on surveys and stuff I'm comfortable sharing in real life as well. I try to be an open book" (P7).

For some participants, selective sharing meant they were only willing to give a certain amount of information during any single HIT. For example, P4 was willing to share either his first name or webcam footage because either one alone did not seem too sensitive, but both together would be too much: "If they had my webcam footage I wouldn't want to give my first name even. If they didn't have my footage I'd give my first name. [No] info that builds on each other. You can have segmented info only, lol" (P4). While participants did not mention costs associated with this PPB, we know from prior work that this can be easily sidestepped by data aggregation across HITs [14].

Reporting HITs. Participants would also report HITs to Amazon for violating TOS, a function built into the MTurk interface. For example, P14 often reported HITs that asked her to connect her social media account: "it's a TOS violation for requesters to ask for that info. If it's egregious I will often report it to Amazon." Participants did not mention any costs to this PPB, possibly since reporting a HIT within the interface is both easy and normative.

Some participants also reported HITs to other Turkers by posting warnings on MTurk community sites: "I do leave reviews on Turkerview, used to on TO [TurkOpticon] when it comes to privacy issues. I'd also post on the forums. Like 'be careful guys, this one wants x and x' " (P4). One participant reported a HIT to an IRB where an academic requester asked highly personal questions for very little pay: "think $1.25 for 60 minutes—I once wrote a nasty email to that requester for taking advantage of people who have experienced trauma, I mean that's just despicable. I copied her IRB too :P" (P14).

Transitions in Privacy Concerns and Practices

As participants learned how to become efficient Turkers, they experienced shifts in their privacy concerns and practices. Becoming an efficient Turker is an arduous process. Most participants expressed dissatisfaction with their initial experience on MTurk, describing a steep learning curve and expressing frustration with how few resources MTurk provided for Turkers: "there is no support system. Nor is there any training. You get accepted and they're sort of like, here you go, get Turking!" (P3). This caused new Turkers large amounts of invisible labor: "in the beginning I could easily spend 8 hours hunting down HITs" (P14).

Crossing the divide from newbie to experienced Turker required its own form of invisible labor, in the form of going beyond MTurk and searching for resources elsewhere on the Internet. Many participants echoed the process described by P4: "I tried on my own first and made next to nothing. Then I started googling and found some forums. Saw what some people made. started trying scripts. It snowballs. So you start out. You have no idea what to do. You see HITs you do them. You get rejected because you don't use TO or read reviews. That's pretty much the cycle nowadays. Then you go on Reddit or the forums and you learn. but everyone screws up at first. I have like 40% of requesters filtered out, because I know better now. But when you start you just try everything. and that's what screws you over."

All participants found the broader MTurk community valuable for learning how to become more efficient, sharing information about HITs and requesters, and increasing their earnings: "Without the help of the other workers on TurkerHub, I would still be convinced that making $20 a day on MTurk is impossible. Now I can make 100 a day on a good day. They've helped me learn so much" (P13). Forums also helped Turkers learn to use scripts to automate finding and accepting lucrative HITs: "I do not know how I did MTurk before I learned how to use [scripts]" (P5).

These improvements in efficiency and earnings empower Turkers, even those who rely on MTurk as a primary source of income, to be more discerning about choosing low-risk HITs and able to engage in PPBs such as returning HITs: "I'm satisfied enough with my earnings at this point that I'm not going to worry about wasting 10 minutes here and there on something I don't submit. [...] But when I was desperate back at the beginning I would do anything for a buck" (P11). Experience with risky HITs also helped most participants become better at staying safe: "I suppose it comes as second nature the more I've studied into this, it's easy for me to look out for what looks too sketchy. I think when I first started I'd be ignorant but now I'm more observant" (P12).

This is not to say that Turkers' privacy concerns necessarily increased over time. Some participants described becoming habituated to providing personal information: "I suppose at the beginning I was more worried about that stuff, but I've done it so often it feels routine now" (P3). In contrast, others became more concerned about privacy over time, feeling they had cumulatively provided a lot of information: "I think I'm [privacy] conscious because I put so much out there if that makes sense" (P14). Privacy scandals in the news could also increase privacy concerns: "After the Cambridge Analytical situation, I have become a bit more cautious with my data. I try to avoid being personally identified" (P13).

5 DISCUSSION

Overall, we found that participants evaluated several privacy-related factors when selecting HITs, that their decision to engage in PPBs was influenced by the benefits and costs of these behaviors, and that their privacy attitudes and practices evolved over time. In the following sections, we discuss what we think these findings tell us about how privacy, power, and invisible labor intersect on MTurk, then summarize the key problems Turkers faced along with ideas for how the MTurk ecosystem could help to address them.

Privacy as Invisible Labor

A useful way to think about our findings is to return to the privacy calculus. Turkers' assessments of the costs and benefits of providing personal information are shaped by power dynamics and economic considerations. They evaluate the privacy-related costs of HITs from multiple perspectives, such as assessing the degree of risk they would be taking on by complying with a privacy-concerning data request. These assessments center on Turkers' perception of the sensitivity and identifiability of the requested information as well as the risks of unauthorized use or inadvertent self-disclosure. Turkers' concerns that their data not be used in unauthorized contexts can be understood by the theory of contextual integrity, which postulates that privacy concerns emerge when information that is disclosed according to the norms of one context is used in another [26].

However, Turkers are limited in their ability to assess contextual integrity risks because of the opacity around requesters' identities and intentions. Moreover, Turkers cannot fully evaluate the privacy risks of HITs up front, since HIT descriptions often do not include the types of information that will be sought—and such requests often come after a Turker has already invested substantial effort. This characterizes one kind of invisible labor for Turkers, who often return HITs when their privacy is threatened halfway into a HIT, losing the effort invested. This likely accounts for at least some of the returned HITs in Hara et al. [12], given

that it was the most commonly reported PPB among our participants.

Turkers' evaluations of benefits were also more complex than merely assessing the HIT's pay rate; for example, some also considered the benefits around contributing to scientific research. We initially suspected that Turkers' economic dependence on the platform would be the main influence on their privacy decisions, which is why we recruited people with varying levels of financial reliance on MTurk. Instead, we found that participants' privacy decisions were influenced more by their ability to find and complete well-paying, low risk tasks while maintaining their target income. Participants found this harder when they were new and inexperienced, and reported completing privacy-invasive HITs out of economic need when they first started out on the platform. Gaining experience and expertise on the platform gave them more power to protect their privacy, even if they depended primarily on MTurk for their income.

Developing the ability to better protect their privacy depended on invisible labor, in the form of researching how to be more efficient. While invisible labor can be costly in terms of hourly earnings, Hara et al. theorize that Turkers learn to minimize unpaid work over time [12]. Our results provide some support for this: the initial invisible labor of becoming efficient on the site allowed Turkers to be more discerning and careful about their privacy in the long-term, reducing wasted effort downstream and making effort that was wasted more economically tolerable.

These observations lead us to conceptualize privacy as a key form of invisible labor on MTurk, which affects Turkers' earnings (e.g., when they abandon a partially completed HIT due to privacy concerns), and puts the onus on Turkers to learn how to recognize risk as well as become more efficient at finding and completing safe and well-paying tasks. Overall, our participants expressed a greater sense of being able to protect their privacy (such as by abandoning risky work and forfeiting payment) once they had put in a substantial amount of invisible labor and learned to maximize their earnings on the site.

Our findings connect to Marwick and boyd's argument that privacy can be a form of privilege, where social position affects one's ability to assert and enforce privacy claims [24]. On MTurk, much of what determines the ability to enact privacy behavior is education and its effects on socio-economic status inside of MTurk: Turkers who engage in the invisible labor of learning about the platform and available resources (such as scripts) are more efficient earners, which provides them with relatively more power to protect themselves. In contrast, new and infrequent Turkers, those who are less technologically savvy or Internet-connected, and those who are less welcome in the forums where Turkers exchange knowledge [11] may be particularly vulnerable to privacy risks. We see a need for future work to examine these subpopulations within MTurk. Overall, we suspect that similar dynamics play out in other digital labor contexts with information and power asymmetries (such as Uber, Figure Eight, and Upwork), and would be interested in seeing parallel analyses of such platforms.

The changes we observed in Turkers' privacy concerns and practices over time also align with models that conceptualize privacy as a dynamic process of informational and boundary management (e.g. [26, 27]), rather than a static, general attitude. Even participants who self-identified as "open books" made many exceptions, aligning with prior findings that contextual factors influence privacy decision-making [28]. Further, there was no dominant narrative about what caused concerns or how they changed over time. As Kittur et al. point out, Turkers are a diverse population with varying reasons for working on MTurk [17]; we believe this diversity is reflected in the range of privacy attitudes and practices we observed. We believe these observations emphasize a more general need to emphasize change, context, and individual differences in privacy research.

Ideas for Reducing Privacy Risks

Our work raises natural questions about what the ecosystems around crowd work platforms such as MTurk might do to reduce privacy risks and empower workers, ideally while also benefiting requesters and the platforms themselves. Below we lay out key barriers that came out of our findings along with proposals toward designs that might mitigate them.

Awareness and Communication. We found that Turkers use the same kinds of information and sources for reasoning about privacy concerns in HITs that they use for other aspects of decision-making, such as pay rate and requester quality. A natural design idea, then, is to make privacy and personal information aspects of HITs more salient on existing community sites. This could include adding explicit privacy-related elements of review forms on sites like Turkopticon beyond the existing flag for indicating TOS concerns, or creating discussion spaces dedicated to privacy and personal information in existing Turker forums. Making privacy information more salient and accessible may reduce the invisible labor involved in returning HITs due to privacy concerns while limiting the ability of insensitive or unscrupulous requesters to collect personal information.

Onboarding and Education. The suggestion above might shift privacy risks toward more vulnerable Turkers, notably newbies, who our participants described as ill-equipped to consider privacy implications. Further, there are serious privacy risks that new Turkers confront right away, including the decision to connect one's MTurk ID to an existing Amazon ID, the use of one's main email address, and scam HITs

that harvest personal information. Helping new workers navigate these risks is important given the continuous inflow of new Turkers [6], but neither MTurk itself nor (to our knowledge) most Turker-created how-to guides highlight considerations about privacy beyond describing the personally identifiable information restrictions in MTurk's TOS.

This is another place where community visibility could go a long way by including privacy as an explicit part of newbie FAQs and related resources. We also see a need to reach Turkers who do not use forums, since our participants were much less empowered to protect themselves before they found external resources. An enterprising and generous individual might choose to, in a requester role, post a penny HIT where the task was to read and review orientation-related materials, to try to help new Turkers more rapidly come up to speed. Even given Difallah et al.'s estimate that there are tens of thousands of new workers every year [6], this could be done on the order of $1,000 a year, likely less. Another path to making this information more widespread is for requesters interested in improving Turkers' efficiency and welfare to include pointers to Turker resources at the end of their HITs.

Transparency and Trust. Turkers operate with limited information about both requesters and HITs, which reduces their ability to assess them, increases invisible labor, and likely reduces trust in both requesters and MTurk. Trust goes a long way in reducing friction in interactions; many of our participants expressed a preference for academic requesters based on the perception that they are more trustworthy, provide information about the task's purpose and the researcher, and are accountable to an IRB.

A privacy by design [2] approach to designing crowd work platforms might require requesters to provide verified information about themselves to build such trust and accountability, not unlike verifications for AirBnB hosts and guests. This might benefit requesters, as providing requester information builds trust and leads Turkers to work harder [22]. Requiring a brief consent form for all HITs could also improve transparency, although Kittur et al. point out that a balance must be struck between providing enough information to allow workers to evaluate a task and creating an informational burden that results in additional invisible labor [17]. Finally, Xia et al. suggested that HITs should list the types of data they collect up front [36]. Our study indicates that such upfront disclosures are particularly important for tasks that request information that Turkers are uncomfortable providing, such as links to social media accounts, and would likely reduce the number of HITs that are returned due to privacy reasons. Requesters who adopt this approach might also avoid being negatively reviewed on MTurk forums.

Supporting Desired Disclosure. While our participants' privacy concerns and practices varied in many ways, some kinds of personal information, such as demographics, were seen as relatively low-risk and reasonable for many tasks. Scripts for automatically entering demographic information are occasionally discussed on MTurk forums, but a lack of standard field names, input widgets, and definitions makes it hard for Turkers to implement them.

It might help both requesters and Turkers if MTurk and Turkers worked together to develop tools for providing relatively benign personal information. Workers could locally store personal information and share it as appropriate based on the context of a given request. Tasks could make requests in machine-readable ways that would support some of the transparency and trust ideas above, making it easier for workers to assess the privacy demands of tasks up front and complete tasks more efficiently. One risk is that this could help Turkers create consistent but false profiles; however, our results suggest that at least for requests perceived as appropriate, Turkers are not malicious and tell relatively few privacy lies. Moreover, semi-automated sharing could reduce data entry errors and privacy lies, improving data quality for requesters.

Protection from Reidentification. Our participants selectively shared information with requesters as a way to segment their data profiles and protect their privacy. This is probably not very effective; as described by Lease et al. [20] and Kandappu et al. [14], even experienced Turkers are unaware that MTurk IDs are often publicly searchable, or that requesters can accumulate profiles of personal information across multiple tasks through the kinds of requests for personal information we just proposed making easier through autofilling.

Unlinking MTurk IDs from Amazon IDs as proposed by Xia et al. [36] addresses the searchability but not the aggregation problem. We propose an alternate idea inspired by "virtual credit cards", short-duration or single-use credit card numbers tied to a customer's actual credit card number. Using virtual cards can reduce the risks of both fraud and profiling by vendors; MTurk could in principle do the same for workers, providing HIT-specific virtual IDs linked to a worker's private MTurk ID. Doing this has real challenges, though. Requesters have legitimate reasons to track behavior across HITs to prevent duplicate survey filling, compute aggregate work quality or preference information, track changes in opinions or activity over time, and so on. Thus, designing the space of allowable queries that balance requester needs with Turker privacy is likely to be hard. Further, we think it is unlikely Amazon would adopt this design, since the idea of a single MTurk ID is probably baked into the platform, and the benefits of such a design would largely

accrue to Turkers rather than being shared between Turkers, requesters, and the platform. However, newer crowdsourcing platforms might use such an idea to help them compete with existing platforms, in part on promises to protect workers' privacy.

6 CONCLUSION

In this study, we identified how privacy considerations affect how and what work gets done on MTurk. We found that Turkers evaluate many factors when judging whether to comply with requests for personal information in HITs, and that while Turkers engage in multiple PPBs, these involve risk, time, and effort. We posit that navigating privacy is a form of invisible labor on MTurk that is exacerbated by the uneven power dynamics on the site.

From the point of view of privacy research, we contribute to the framing of privacy work as a kind of invisible labor, where the ability to protect one's privacy is realized through the effort Turkers put in to empower themselves to be more effective workers. We suspect that invisible labor may be a fruitful lens to think about privacy behavior in other contexts as well. Our findings also support the value of considering privacy as a fluid, personal, and contextual process, and provide a case study of how studying the privacy needs of particularly vulnerable populations can reveal universal privacy concerns and issues.

Toward crowdsourcing, we provide additional insight into Turkers' work practices, highlighting the considerations and invisible labor involved in protecting themselves against privacy risks. Our work also reveals the ironic side effect that committed Turkers working hard to empower themselves may inadvertently increase the vulnerability of new Turkers. Our findings also help to understand why HITs are returned, partially explaining a major inefficiency for both Turkers and requesters, and provide the basis for a number of design ideas to mitigate negative privacy-related effects in both MTurk and digital labor markets more generally.

ACKNOWLEDGMENTS

This research was supported by a seed grant from Cornell University's Center for the Study of Inequality to the first author. It was conducted while the second author was serving at the National Science Foundation and does not necessarily reflect the views of the NSF.

REFERENCES

[1] Ali Alkhatib, Michael S Bernstein, and Margaret Levi. 2017. Examining Crowd Work and Gig Work through the Historical Lens of Piecework. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*. ACM, 4599–4616.

[2] Ann Cavoukian. 2011. Privacy by Design in Law, Policy and Practice. *A white paper for regulators, decision-makers and policy-makers* (2011).

[3] Kathy Charmaz. 2001. *Qualitative Interviewing and Grounded Theory Analysis*. 675–694.

[4] Lydia B Chilton, John J Horton, Robert C Miller, and Shiri Azenkot. 2010. Task Search in a Human Computation Market. In *Proceedings of the ACM SIGKDD Workshop on Human Computation*. ACM, 1–9.

[5] Mary J Culnan and Pamela K Armstrong. 1999. Information Privacy Concerns, Procedural Fairness, and Impersonal Trust: An Empirical Investigation. *Organization Science* 10, 1 (1999), 104–115.

[6] Djellel Difallah, Elena Filatova, and Panos Ipeirotis. 2018. Demographics and dynamics of Mechanical Turk Workers. In *Proceedings of the Eleventh ACM International Conference on Web Search and Data Mining*. ACM, 135–143.

[7] Djellel Eddine Difallah, Michele Catasta, Gianluca Demartini, Panagiotis G Ipeirotis, and Philippe Cudré-Mauroux. 2015. The Dynamics of Micro-task Crowdsourcing: The Case of Amazon MTurk. In *Proceedings of the 24th International Conference on World Wide Web*. International World Wide Web Conferences Steering Committee, 238–247.

[8] Alek Felstiner. 2011. Working the Crowd: Employment and Labor Law in the Crowdsourcing Industry. *Berkeley J. Emp. & Lab. L.* 32 (2011), 143.

[9] Ujwal Gadiraju, Ricardo Kawase, and Stefan Dietze. 2014. A Taxonomy of Microtasks on the Web. In *Proceedings of the 25th ACM Conference on Hypertext and Social Media*. ACM, 218–223.

[10] Barney G Glaser and Anselm L Strauss. 1967. The Discovery of Grounded Theory: Strategies for Qualitative Theory. *New Brunswick: Aldine Transaction* (1967).

[11] Mary L Gray, Siddharth Suri, Syed Shoaib Ali, and Deepti Kulkarni. 2016. The Crowd is a Collaborative Network. In *Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing*. ACM, 134–147.

[12] Kotaro Hara, Abigail Adams, Kristy Milland, Saiph Savage, Chris Callison-Burch, and Jeffrey P Bigham. 2018. A Data-Driven Analysis of Workers' Earnings on Amazon Mechanical Turk. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. ACM, 449.

[13] Lilly C Irani and M Silberman. 2013. Turkopticon: Interrupting Worker Invisibility in Amazon Mechanical Turk. In *Proceedings of the 2013 CHI Conference on Human Factors in Computing Systems*. ACM, 611–620.

[14] Thivya Kandappu, Vijay Sivaraman, Arik Friedman, and Roksana Boreli. 2013. Exposing and Mitigating Privacy Loss in Crowdsourced Survey Platforms. In *Proceedings of the 2013 Workshop on Student Workhop (CoNEXT Student Workhop '13)*. ACM, New York, NY, USA, 13–16.

[15] Ruogu Kang, Stephanie Brown, Laura Dabbish, and Sara Kiesler. 2014. Privacy Attitudes of Mechanical Turk Workers and the US Public. In *Symposium on Usable Privacy and Security (SOUPS)*, Vol. 4. 1.

[16] Nicolas Kaufmann, Thimo Schulze, and Daniel Veit. 2011. More than Fun and Money: Worker Motivation in Crowdsourcing – A Study on Mechanical Turk. In *AMCIS*, Vol. 11. 1–11.

[17] Aniket Kittur, Jeffrey V Nickerson, Michael Bernstein, Elizabeth Gerber, Aaron Shaw, John Zimmerman, Matt Lease, and John Horton. 2013. The Future of Crowd Work. In *Proceedings of the 2013 Conference on Computer-Supported Cooperative Work*. ACM, 1301–1318.

[18] Walter S Lasecki, Jeffrey M Rzeszotarski, Adam Marcus, and Jeffrey P Bigham. 2015. The Effects of Sequence and Delay on Crowd Work. In *Proceedings of the 2015 ACM Conference on Human Factors in Computing Systems*. ACM, 1375–1378.

[19] Walter S Lasecki, Jaime Teevan, and Ece Kamar. 2014. Information extraction and manipulation threats in crowd-powered systems. In *Proceedings of the 17th ACM Conference on Computer-Supported Cooperative Work & Social Computing*. ACM, 248–256.

[20] Matthew Lease, Jessica Hullman, Jeffrey Bigham, Michael Bernstein, Juho Kim, Walter Lasecki, Saeideh Bakhshi, Tanushree Mitra, and

Robert Miller. 2013. Mechanical Turk is not Anonymous. (2013).

[21] J Lofland, D Snow, L Anderson, and LH Lofland. 2006. *Analyzing Social Settings: A Guide to Qualitative Observation and Analysis*. Wadsworth/Thomson Learning.

[22] Jennifer Marlow and Laura A Dabbish. 2014. Who's the Boss?: Requester Transparency and Motivation in a Microtask Marketplace. In *CHI'14 Extended Abstracts on Human Factors in Computing Systems*. ACM, 2533–2538.

[23] David Martin, Benjamin V Hanrahan, Jacki O'Neill, and Neha Gupta. 2014. Being a Turker. In *Proceedings of the 17th ACM Conference on Computer-Supported Cooperative Work & Social Computing*. ACM, 224–235.

[24] Alice E Marwick and danah boyd. 2018. Understanding Privacy at the Margins – Introduction. *International Journal of Communication* 12 (2018), 9.

[25] Brian McInnis, Dan Cosley, Chaebong Nam, and Gilly Leshed. 2016. Taking a HIT: Designing around Rejection, Mistrust, Risk, and Workers' Experiences in Amazon Mechanical Turk. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*. ACM, 2271–2282.

[26] Helen Nissenbaum. 2004. Privacy as Contextual Integrity. *Washington Law Review* 79 (2004), 119.

[27] Sandra Petronio. 2012. *Boundaries of Privacy: Dialectics of Disclosure*. SUNY Press.

[28] Shruti Sannon, Natalya N Bazarova, and Dan Cosley. 2018. Privacy Lies: Understanding How, When, and Why People Lie to Protect their Privacy in Multiple Online Contexts. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. ACM, 52.

[29] Shruti Sannon and Dan Cosley. 2018. It was a shady HIT: Navigating Work-Related Privacy Concerns on MTurk. In *Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems*. ACM, LBW507.

[30] Trebor Scholz. 2012. *Digital labor: The Internet as Playground and Factory*. Routledge.

[31] Aaron Smith. 2016. Gig work, online selling and home sharing. *Pew Research Center* 17 (2016).

[32] Jai-Yeol Son and Sung S Kim. 2008. Internet Users' Information Privacy-protective Responses: A Taxonomy and a Nomological Model. *MIS Quarterly* (2008), 503–529.

[33] Susan Leigh Star and Anselm Strauss. 1999. Layers of Silence, Arenas of Voice: The Ecology of Visible and Invisible Work. *Computer-Supported Cooperative Work (CSCW)* 8, 1-2 (1999), 9–30.

[34] Latanya Sweeney. 2000. Simple demographics often identify people uniquely. *Carnegie Mellon, Data Privacy Working Paper 3. Pittsburgh* (2000), 1–34.

[35] Amazon Mechanical Turk. 2018. Acceptable Use Policy. (2018). https://www.mturk.com/worker/acceptable-use-policy

[36] Huichuan Xia, Yang Wang, Yun Huang, and Anuj Shah. 2017. "Our privacy needs to be protected at all costs": Crowd workers' privacy experiences on Amazon Mechanical Turk. *Proc. ACM Hum.-Comput. Interact* 1 (2017).

[37] Jie Yang, Carlo van der Valk, Tobias Hossfeld, Judith Redi, and Alessandro Bozzon. 2018. *How do Crowdworker Communities and Microtask Markets Influence Each Other? A Data-driven Study on Amazon Mechanical Turk*. Technical Report. Université de Fribourg.

[38] Ming Yin, Mary L Gray, Siddharth Suri, and Jennifer Wortman Vaughan. 2016. The Communication Network Within the Crowd. In *Proceedings of the 25th International Conference on World Wide Web*. International World Wide Web Conferences Steering Committee, 1293–1303.