

# Investigating the Accessibility of Crowdwork Tasks on Mechanical Turk

Stephen Uzor  
Department of Engineering  
University of Cambridge  
Cambridge, United Kingdom  
scu20@cam.ac.uk

John J. Dudley  
Department of Engineering  
University of Cambridge  
Cambridge, United Kingdom  
jjd50@cam.ac.uk

Jason T. Jacques  
Department of Engineering  
University of Cambridge  
Cambridge, United Kingdom  
jtj21@cam.ac.uk

Per Ola Kristensson  
Department of Engineering  
University of Cambridge  
Cambridge, United Kingdom  
pok21@cam.ac.uk

## ABSTRACT

Crowdwork can enable invaluable opportunities for people with disabilities, not least the work flexibility and the ability to work from home, especially during the current Covid-19 pandemic. This paper investigates how engagement in crowdwork tasks is affected by individual disabilities and the resulting implications for HCI. We first surveyed 1,000 Amazon Mechanical Turk (AMT) workers to identify demographics of crowdworkers who identify as having various disabilities within the AMT ecosystem—including vision, hearing, cognition/mental, mobility, reading and motor impairments. Through a second focused survey and follow-up interviews, we provide insights into how respondents cope with crowdwork tasks. We found that standard task factors, such as task completion time and presentation, often do not account for the needs of users with disabilities, resulting in anxiety and a feeling of depression on occasion. We discuss how to alleviate barriers to enable effective interaction for crowdworkers with disabilities.

## CCS CONCEPTS

• **Human-centered computing** → **Empirical studies in accessibility**; • **Information systems** → **Crowdsourcing**.

## KEYWORDS

crowdsourcing, crowdwork, AMT, MTurk, accessibility, disability

## ACM Reference Format:

Stephen Uzor, Jason T. Jacques, John J. Dudley, and Per Ola Kristensson. 2021. Investigating the Accessibility of Crowdwork Tasks on Mechanical Turk. In *CHI Conference on Human Factors in Computing Systems (CHI '21)*, May 8–13, 2021, Yokohama, Japan. ACM, New York, NY, USA, 14 pages. <https://doi.org/10.1145/3411764.3445291>



This work is licensed under a Creative Commons Attribution International 4.0 License.

*CHI '21, May 8–13, 2021, Yokohama, Japan*  
© 2021 Copyright held by the owner/author(s).  
ACM ISBN 978-1-4503-8096-6/21/05.  
<https://doi.org/10.1145/3411764.3445291>

## 1 INTRODUCTION

Crowdwork is a relatively new form of digital labor where crowdworkers complete microtasks posted by a ‘requester’ in order to earn income [8, 61]. Given the flexibility involved in crowdwork and the lack of a need to commute to work, it is unsurprising that crowdwork offers several advantages for work, which in most cases can be ideal for people with disabilities [61].

Amazon Mechanical Turk (AMT) is a popular crowdwork platform, launched by Amazon in 2005, which has seen increasing use as a tool for research, including in social science [48], education [20], natural language processing (NLP) [57] image labelling [53, 58] and more [35, 38]. Using crowdwork for user research raises a number of issues, such as a lack of experimental controls, a need to ensure users invest effort in the tasks and a requirement and desire to maintain high research ethics standards [12, 23]. Nonetheless, a series of studies have demonstrated that the results from crowdwork studies are practically indistinguishable from those that use more ‘traditional’ recruitment channels [11, 19, 20, 28, 56]. Other notable crowdwork platforms include Clickworker and Microworkers, although there are numerous such platforms available online.

Common microtasks (known as Human Intelligence Tasks, or HITs) performed on AMT include surveys, content generation, audio and video transcription, language translation, information finding and verification and validation. Users can engage in crowdwork using a relatively simple and inexpensive computer connected to the Internet, which is another factor explaining why this type of work has become a viable option for either primary or secondary work [8, 61].

Despite the increasing interest in crowdwork by researchers and users alike, the literature suggests that crowdwork presents several accessibility issues [10, 37, 39, 60, 61], which can impact successful engagement in this type of work—examples of which include user interface issues [10], language barriers [37], wages [61], time limits [8, 61] and unclear instructions [8]. Despite this body of literature, questions remain unanswered as to how the above issues affect people with a range of disabilities that are most likely to affect computer use. Further, it is important to identify how people with a range of disabilities approach or avoid crowdwork tasks, what role their impairment plays in the successful engagement in crowdwork and their opinions on how these issues may be alleviated.

In this paper, we first conducted two surveys to understand how AMT workers who identify as having a disability cope with performing HITs (microtasks). We explored a variety of the most common disability categories [16], including: cognition/mental, hearing, physical, mobility, visual and motor/dexterity impairments. Out of the 1,000 participants that responded to the first survey, 626 reported having a disability. We invited a subset of the participants in the first survey (100 from each disability category) to take part in the second survey, which specifically focused on the individual disability categories. Following both surveys, we conducted further interviews with 5 participants. We present demographic information on those users who identify as having a disability and discuss the findings in light of previous work. Furthermore, we discuss how identified barriers to accessibility can be reduced or alleviated to foster effective crowdwork. Given the increasing adoption of crowdwork as a permanent or supplemental source of income for users, the implications of such an investigation for HCI cannot be overstated. Hence, to the best of our knowledge, our work makes the following contributions to the wider HCI literature:

- (1) An investigation into the demographics of crowdworkers, in the US, who identify as having a disability and use the AMT platform.
- (2) Insights into the factors that hinder the effective completion of crowdwork tasks by people with a variety of difficulties and impairments that are most likely to hinder computer use.
- (3) A discussion on how identified accessibility issues can be alleviated in order to encourage effective participation in crowdwork.

## 2 BACKGROUND

### 2.1 Crowdwork

Crowdsourcing is a relatively new technology for outsourcing work [33, 53, 56] in the form of microtasks to people who are generally known as crowdworkers [3]. Hensel et al. [29] define crowdworking as the subset of crowdsourcing that involves paid work. Through crowdwork platforms, microtasks (a unit of a task on a crowdwork platform) can be available to thousands of workers from diverse backgrounds and communities. This affords potential benefits for research [1, 12, 27, 35, 38, 41, 58], such as the ability to sample a relatively large number of users from the population, high statistical power and generalizability of results [23, 51]. Nevertheless, in certain cases such crowdwork platforms may be unsuitable for HCI-related experiments [30] due to a lack of experimental controls, means of ensuring user investment in the task and ethics [12, 23]. Popular crowdwork platforms [23] include Clickworker [15] and Mechanical Turk [2], although there are numerous available alternatives. We focus on Mechanical Turk in this paper, not only because of the platform's popularity [25, 41] and widely known brand association [41], but also due to the fact that it has been extensively studied in the literature [25, 28, 31, 36, 44], including in HCI (e.g. [25, 53]), thereby facilitating a potential for comparison between studies.

**2.1.1 Amazon Mechanical Turk (AMT) and Terminology.** Mechanical Turk (also known as AMT or MTurk) is a crowdsourcing platform developed by Amazon in 2005 that allows microtasks to be completed by crowdworkers [46, 48, 53]. A person who posts a microtask on AMT is called a requester, and the crowdworkers are known simply as workers, or 'Turkers' (when specifically referring to AMT workers). In AMT, a microtask is referred to as a Human Intelligence Task (HIT), since these tasks are mostly suitable for human processing, rather than a machine. We use the terminology defined in this section in the rest of the paper.

AMT normally requires certain qualifications from workers in order to access and complete HITs on the platform [25]. Such qualifications can range from having a certain number of previous HITs completed to having a 'masters' qualification gained by consistently completing HITs of a certain type from a variety of requesters with a high degree of accuracy. On the completion of HITs, requesters evaluate the HITs and can either approve or reject them. Approval results in payment, whereas a rejection does not. If a HIT is accepted but not completed, it is returned.

Monetary rewards for completing HITs on AMT can be as little \$0.01, with a median pay rate at \$3.13 [25]. Given that this is far below the minimum wage rate in the US, it can lead to unsatisfactory and stressful working conditions [7, 25]. Certain studies have found that income generation is the primary reason many workers engage in crowdwork [8, 25, 45]. Therefore, low compensation presents a problem to many workers [26]. Consequently, workers often resort to strategies, such as using scripts and community support to identify suitable HITs, including those with relatively higher payments [25].

### 2.2 Demographics of Crowdworkers on AMT

Numerous studies have been conducted to investigate the demographics of AMT workers [17, 53]. Ross et al. [53] estimated more than 400,000 registered workers, with some research placing the worker count at 500,000 [12, 40]. It is not known how many of the users are active [12], although US workers were estimated at 15,000 at any given time [21, 59]. However, in the interest of the research community, previous work revealed that only 7,300 participate in research studies [59], although Difallah et al. [17] estimate this number to be closer to 2,450 as at 2018. Prior work [42] also suggests that even though recruiting workers with the *masters* qualification provides some quality assurance, there is no marked difference in quality of results between more experienced workers and their less experienced peers.

One question that arises at this point is: *Who are AMT workers?* An informal 2010 investigation [34] suggested that the AMT demographics primarily included workers based in the US who are relatively young, female, educated and with moderately high incomes. Formal studies confirmed this demographic [17, 49] but highlight an increase in international workers, mostly from India [17, 37].

The literature highlights efforts to understand crowdworker demographics with regard to users identifying as having a disability. For example, Zyskowski et al. [61] explore the challenges associated with microtask employment for people with disabilities. They performed a survey with over 600 participants and carried out follow

up interviews with 13 respondents. Zyskowski et al. [61] found that a reasonable number of people with disabilities engage in crowdwork. Consequently, they outline a number of avenues for further research, including understanding: a) the level of representation of people with different specific disabilities in crowdwork; and b) the ever-changing demographics (e.g. the elderly) of people with disabilities and the potential dynamics for support within this community. Zyskowski et al.'s work [61] focuses on attitudes of people with disabilities towards crowdwork, regardless of whether or not they actually do engage, or have previously engaged, in crowdwork. In their study, a relatively small subset of participants used AMT. Our work, while similar to [61], differs mainly in the sense that we focus specifically on AMT and aim to understand the demographics of users with different disabilities and how these disabilities affect their engagement in crowdwork on AMT.

### 2.3 Accessibility in Crowdwork

In an attempt to evaluate the accessibility of AMT from a strict user interface (UI) standpoint, Calvo et al. [10] conducted a heuristic analysis of AMT using the Web Content Accessibility Guidelines (WCAG) 2.0 [52]—a set of guidelines for making websites useful and more accessible for users. Calvo et al. [10] found a variety of potential accessibility challenges with the AMT interface (in 2014), which could present barriers to use for those with visual, audio, cognitive and physical disabilities. Although the work by Calvo et al. [10] provided useful insights into how the AMT interface could be improved, there are a number of key opportunities for further work in the HCI domain. First, it is important to understand how AMT crowdworkers perceive accessibility issues and how these issues hinder their effective use of the platform. Second, although heuristic evaluations are a valid commonly used method in HCI, studies have suggested that expert audits of the WCAG guidelines do not meet the threshold of 80% agreement necessary to give a definitive answer on whether or not guidelines have been met [4, 5]. Therefore, it is useful to gain insights into users' experiences with crowdwork platforms, by engaging in conversations with users.

Researchers have conducted such investigations to improve the accessibility of crowdwork for users with low information and communication technology (ICT) literacy. For instance, in 2010, Khanna et al. [37] conducted a study to evaluate the usability of AMT interfaces for low-income workers in India. They proposed design recommendations based on the fact that the interfaces for HITs were often too complicated for these users, especially regarding the English language. Furthermore, a 2013 study by Kobayashi et al. [39] developed a question-answer card interface to allow low ICT-literate elderly users in Japan to participate in crowdwork with minimal effort.

Certain disabilities are often associated with aging, and prior work has investigated barriers to effective participation in crowdwork in older adults. For instance, Brewer et al [8] report that older adults are more likely to engage in crowdwork, if certain accessibility issues, such as 'unfair' time limits and unclear instructions, are revised. Zyskowski et al. [61] also highlight accessibility issues involving people with disabilities who engage in crowdwork, and they suggest strategies for assisting such users, who may be unaware of crowdwork, in finding opportunities in this space.

## 3 RESEARCH OBJECTIVES

Most of the important works in the literature have investigated general accessibility in crowdwork, which provides an essential starting point for our work. To address the gaps in the literature, we set out to achieve the objectives below. We focus on AMT workers based in the US in order to compare demographics with existing research in this area (e.g. [61]) and the US Census Bureau [9].

- (1) Investigate the demographics of users who identify as having a disability and who regularly engage in crowdwork on AMT.
- (2) Identify how these users are limited by their individual disabilities and how accessibility issues identified in the literature may be exacerbated given specific disabilities, such as cognition/mental, vision, hearing, reading and motor/dexterity.
- (3) Discuss design implications for the AMT platform and HITs, based on our research findings, in order to enable effective use of the platform by people with disabilities.

## 4 METHOD

Using a mixed-methods approach, we conducted two surveys and follow-up interviews with survey respondents. Both surveys were administered over a three-week period in the summer of 2020 and were presented to participants as HITs on the AMT platform. The first survey was a general survey to gain insights into user demographics, including any existing disabilities. The title of the HIT indicated that it was a *survey on accessibility in crowdwork*. We expected this HIT to attract people who identify as having a disability, in addition to non-disabled people who just wanted to complete the HIT. We then targeted participants who responded in the affirmative in the first survey as having at least one disability and invited them privately to participate in the second survey, which was only made available to eligible respondents from Survey 1. Note that there was no initial indication in Survey 1 that the participants would be completing another HIT for a further survey. The second survey was a focused survey to gain insights into barriers to the successful completion of HITs based on users' existing disabilities. A secondary benefit for conducting a second focused survey was to eliminate potential bias due to misrepresentation, sometimes done by crowdworkers in order to qualify for a HIT [55] in the competitive crowdwork sphere.

We performed consistency checks on the survey results to identify unusual patterns of answers and to confirm that there was a set of sufficiently diverse answers [3, 23]. Open-ended questions from the focused survey and transcribed interview responses were analyzed using a hybrid process of inductive and deductive thematic analysis methods to develop concepts, themes and interpretations based on the data [13, 18]. We first used a deductive approach to identify data patterns that were relevant to the disability themes of interest to our work. We then used open coding [47] to explore the qualitative data to allow for the discovery of emergent themes, which we had previously not identified [61].

Payments for most HITs on AMT are typically low (average of about \$2/hr [25]). Our surveys were designed to meet the US federal minimum wage (\$7.25/hr) and were presented as separate HITs with \$0.25 and \$0.20 rewards for Surveys 1 and 2 respectively. Ethics approval was granted for the study by the authors' university research ethics committee.

## 4.1 Survey 1: General Survey

The questions in this survey were split into two categories: a) demographic information, such as: age, gender, race, employment status and education level; and b) whether participants identified as having certain disabilities. The questions on the latter were adapted from the US Census Bureau [6] with some requiring binary answers (yes/no) and others requiring scale-based answers to determine the seriousness of the issue—see example of these questions below:

- Vision: *Do you have serious difficulty seeing, even when wearing glasses? (yes, no)*
- Hearing: *Do you have serious difficulty hearing? (yes, no)*
- Cognition/Mental: *Because of a physical, mental or emotional condition, do you have serious difficulty concentrating, remembering things or making decisions? (yes, no)*
- Reading: *Do you have difficulty reading text on the screen, without an assistive device (for example, screen reader / magnifier)? (yes, no)*
- Mobility: *Do you have serious difficulty walking or climbing stairs? (yes, no)*
- Motor/Dexterity: *How well can you move the cursor on the screen using, for example, a mouse, trackball or touchpad? (1: with extreme difficulty, 2: with major difficulty, 3: with some difficulty, 4: with minor difficulty, 5: with no difficulty)*

To target those participants who may experience accessibility issues when using the AMT platform, we invited participants who responded ‘yes’ to the binary questions and who responded with 1–4 in the scale-based questions to take part in Survey 2. Survey 1 was completed by 1,000 participants.

## 4.2 Survey 2: Focused Survey

Survey 2 was presented as a multi-dimensional survey HIT on AMT, with one section for general questions and individual dimensions for each disability category. The general section and one dimension of the survey were completed by 100 eligible participants, on a first-come-first-served basis, from each disability category in Survey 1. For instance, 100 participants who reported having visual impairments in Survey 1 completed Survey 2’s general questions section and the visual disability component. Therefore there were 600 total responses for all disability categories. Survey 2 was tailored in this way so that participants only needed to complete a survey that was specific to their identified disability. In total, 600 responses were received for all disability categories. General questions were focused the following topics:

- (1) *Reasons for using AMT.*
- (2) *Whether earnings from AMT are sufficient for daily living.*
- (3) *What AMT HITs are difficult to perform, given the participant’s identified disability.*
- (4) *Whether performance of AMT HITs induces, or is affected by, anxiety.*
- (5) *Whether performance of AMT HITs induces, or is affected by, a feeling of depression.*

In Survey 2, questions were tailored to identify difficulties completing HITs given the participants’ identified disabilities—these included binary, scale and open-ended answers depending on the question (see Table 1). Certain questions were adapted from the

US Census Bureau [6, 9]. In all of the categories, we asked the participants how their use of AMT was affected by their disability (open-ended) and to offer suggestions for whether, and how, any identified issues could be alleviated.

## 4.3 Interviews

Following the completion of both surveys, we invited participants from Survey 2 to take part in an online interview (using a HIT), in order to capture more detailed information on the barriers to completing HITs on the AMT platform. The five participants who responded to the interview call were paid \$5 in remuneration, which lasted for 30 minutes. Interview questions were focused on the following points: a) difficulties completing HITs; b) income; c) work goals; d) user interface; and e) opportunities for improving accessibility. Example questions that we asked are as follows:

- *“What role does the ability to work from home play in performing crowdwork on AMT?”*
- *“What difficulties do you have when completing HITs due to any existing impairment(s)?”*
- *“What HITs do you complete and/or avoid as a result of your existing impairment(s)?”*
- *“What can Amazon or third party requesters do to help you perform this HIT or alleviate accessibility issues?”*

## 5 SURVEY 1 FINDINGS

Based on the results, 626/1,000 (62.6%) of the survey respondents reported having at least one disability—Table 2 shows the percentage of respondents per disability dimension according to the survey. Note that certain people reported having more than one disability. Just over a third (38%) of all respondents reported that they use an assistive device, for example a magnifying tool, sticky keys or text-to-speech reader; this finding is in line with previous work which found that 39.2% of respondents use assistive devices to participate in crowdwork [61]. In the interest of the current research, we focus on these respondents who identified as having at least one disability or at least a minor difficulty with dexterity (mouse and keyboard).

### 5.1 Age Distribution

The distribution of the participants’ ages (see Table 3) was skewed towards younger users. This is not surprising since previous work shows a relatively young demographic on AMT [34, 53]. However this result is interesting in the sense that the average age of respondents who reported having a disability is identical to that of the general population, despite the increased prevalence of disabilities in older age [61]. Our findings are contrary to that of Zyskowski et al. [61], where the demographic was relatively skewed towards an older population. However, Zyskowski et al. [61] surveyed individuals using the Cint service [14], and other avenues, who may or may not engage in crowdwork; whereas our participants were all recruited through AMT, and are therefore, more representative of users who identify as disabled and engage in crowdwork.

### 5.2 Gender

Gender was slightly skewed towards males in both the general respondents and those who reported having a disability—see Table 4.

**Table 1: Example questions from individual disability dimensions in Survey 2**

Dimension	Question
Cognition/Mental	<i>Do you frequently get confused about what day or time it is or where you are? (Binary)</i>
Vision	<i>Do you find the AMT interface difficult to navigate? (Binary &amp; Open-ended)</i>
Hearing	<i>Can you understand loud speech in a quiet room? (Binary)</i> <i>Do you have any difficulties completing an audio transcription HIT? (Binary &amp; Open-ended)</i>
Reading	<i>Are there any HITs that you find difficult due to reading limitations (Binary)</i>
Motor/Dexterity	<i>Can you tie a bow in laces or string without difficulty? (Binary)</i> <i>How easy/ difficult do you find plugging in a USB cable? (Likert scale 1-5)</i>

**Table 2: Number of respondents who reported having a disability in each of the explored dimensions. Note: in the motor impairment/ dexterity scales, 1 = ‘extreme difficulty’; 2 = ‘major difficulty’; 3 = ‘some difficulty’; 4 = ‘minor difficulty’; 5 = ‘no difficulty’**

Disability	% of Respondents	Mean (SD) of difficulty rating
Visual impairment	27	–
Hearing impairment	27	–
Cognitive/mental impairment	30	–
Motor/dexterity impairment (with a mouse)	51	3.5 (1.02)
Motor/dexterity impairment (with a keyboard)	49	3.6 (0.99)
Mobility impairment	28	–
Reading impairment	27	–

**Table 3: Age distribution in Survey 1 (General survey)**

Age	All respondents (%)	Respondents with 1+ disability (%)
Mean	34.6	34.8
Median	32	31
SD	10.2	10.8

The literature describes a demographic skewed in favour of females as of 2010 [34, 53]. Our findings indicate that more males are now using AMT and that participants who identify as having a disability are representative of the general sample.

**Table 4: Distribution of Gender in Survey 1 (General survey)**

Gender	All respondents (%)	Respondents with 1+ disability (%)
Male	57.8	55.4
Female	42	44.4
Other	0.2	0.2

### 5.3 Race

Participant demographics were heavily skewed towards white users—see Table 5. This is, perhaps, unsurprising given that the distribution of race, as found in this study, is similar to that reported by the

**Table 5: Distribution of Race in Survey 1 (General survey)**

Race	All respondents (%)	Respondents with 1+ disability (%)
White	75.7	72.6
Black	13.9	17.2
Asian	3.7	3.7
Native	3.7	3.1
Hispanic	2.2	1.68
Mixed	0.8	0.3
Other	0.3	0.37
Undeclared	0.3	0.19

US Census Bureau, with the exception of the Hispanic population, which represents about 18.5% of the US population [9].

### 5.4 Education Level

A majority of the respondents (83%) reported that they have a bachelor’s or higher degree, highlighting a relatively literate demographic—see Table 6. Similar to Age, our findings support the literature, which shows a relatively highly-educated demographic among AMT workers in the US [34, 53].

### 5.5 Employment

Three quarters of the respondents were permanently employed in some capacity (working for less, or more, than 35 hours a week), while 21% or respondents were self-employed—see Table 7. Almost half of the respondents engaged in work for less than 35 hours a

**Table 6: Distribution of Education in Survey 1 (General survey)**

Education Level	All respondents (%)	Respondents with 1+ disability (%)
Bachelor's degree or higher	77.5	83.4
College	14.7	12.1
High school	7.3	3.9
Other	0.3	0.37
Primary school	0.1	0
None	0.1	0

**Table 7: Distribution of Employment in Survey 1 (General survey)**

Employment type	All respondents (%)	Respondents with 1+ disability (%)
Employed <35 hours	40.5	44.2
Employed 35+ hours	38	31.9
Self employed	16.9	20.5
Unemployed	2.2	0.9
Other	1.5	0.9
Unable to work	0.8	1.3
Retired	0.1	0

**Table 8: Total time per week spent completing HITs on AMT in Survey 1 (General survey)**

Number of hours	All respondents (%)	Respondents with 1+ disability (%)
Less than 5	6.1	6.0
5 - 10	29.8	35.3
11 - 20	28.5	27.8
21 - 35	21.1	17.9
More than 35	14.5	12.9

week. Given that almost 80% of the respondents reported that they are permanently employed, the findings suggest that most of the respondents do not engage in crowdwork as a primary means of employment.

## 5.6 Time Spent Working on AMT

The amount of time that the respondents spent working on AMT was varied, with most people spending 5–9 (35%) or 10–19 (28%) hours per week and only 13% of respondents spending more than 35 hours a week completing tasks on AMT—see Table 8. This result is not surprising given that, according to the findings on employment, a majority of the respondents reported being permanently employed.

**Table 9: Most common HITs completed on AMT in Survey 1 (General survey)**

Employment Type	All respondents (%)	Respondents with 1+ disability (%)
Survey	81.9	69.4
Information finding	42.6	39.9
Verification and validation	40.1	41.4
Interpretation and analysis	38.9	29.9
Content creation	21.5	15.5
Content access	18.8	9.5

## 5.7 Commonly Completed Tasks on AMT in Survey 1

Table 9 highlights the ratios of the most commonly completed tasks (based on the taxonomy of microtasks defined in [22]) on AMT according to our survey. Note that most participants reported that they performed more than one task. For instance, 69% and 41% of respondents reported that they regularly completed *surveys* and *verification and validation* tasks respectively.

## 5.8 Location

Every state in the US, except Maine, Vermont and Wisconsin was represented in the survey among respondents who reported at least one disability. Of the respondents who were geo-located to state level, the highest numbers were from New York (15.2%), California (13.9%), Texas (7.4%), Florida (4.5%) and Illinois (3.8%).

## 6 SURVEY 2 FINDINGS

In this section we first report on the general questions that were available to all of the respondents. Thereafter we discuss the results from the individual survey disabilities that are most likely to affect crowdwork or computer use [16]—vision, hearing, cognition/mental, motor/dexterity and reading.

### 6.1 Working on AMT

Respondents were asked to state their primary reason(s) for using the AMT platform. Of the choices presented, 74% of all Survey 2 respondents stated that they primarily used AMT because it allowed them to work from home; 33% used AMT to earn a living (with no additional income); 28% used the income from AMT to supplement their existing income from employment, and 23% of respondents reported that they used AMT because they enjoyed completing the HITs.

### 6.2 Difficulties Experienced when Completing HITs on AMT

Respondents were asked to comment on any difficulties they have completing HITs on AMT, to which 43% reported that the allocated time for certain HITs was often not enough. There were various reasons why the participants were not satisfied with the allocated

time to complete these microtasks. For instance, they felt that HITs often required them to think carefully about responses, however, the time given did not allow them to do this effectively.

*“Sometimes the instructions in game type HITs are too complicated for me to understand as the language isn’t plain or simple enough. If I finally understand it, by the time I do there’s too little time remaining in the hit [...] and at some point I end up returning it...”* P254 (Hearing impaired user.)

*“Some long surveys need to be understood and to give wise answers. So it is takes time to complete in the allocated time”* P132 (Motor impaired user.)

Time was not the only factor that presented difficulties to the participants when completing HITs. For instance, the above comment by P254 highlights difficulties due to instructions that are perceived as complicated. P600 also felt that difficulties can arise when instructions are lacking sufficient information for the successful completion of the task.

*“Some [HITs] need more time to perform but they will give one or two [HITs] like collect data from websites. Sometimes they don’t even give links to collect data.”* P600 (Visually impaired user)

### 6.3 Effects of Perceived Anxiety and Depression on HIT Completion

We asked the respondents whether they found task completion difficult for any HITs due to perceived anxiety or depression; 39% and 37% of all Survey 2 respondents responded in the affirmative for perceived anxiety and depression respectively. Note that this does not mean that they had been diagnosed for either condition (e.g. clinical depression), but that they experienced (perceived) one or both conditions as a consequence of performing crowdwork.

**6.3.1 Perceived Anxiety.** According to P27, the nature of the HIT could induce anxiety in them if it encouraged them to recall a traumatic event.

*“Sometimes tasks that delve into psychological trauma are too much for me given my background.”* P27 (Cognitively impaired user.)

Other participants stated that anxiety can arise when they feel that their worker reputation is at stake, especially since this reputation can be critical to their reliance on crowdwork for income.

*“Most [HITs] just induce anxiety these days because you could reject me for no reason and I depend on this site for money.”* P61 (Cognitively impaired user.)

*“[...]you can have one requester give you some rejections and your approval rating can get low. I like to keep mine above 99.50.”* P400 (Motor impaired user)

In addition to presenting other difficulties as mentioned in Section 6.2, the allocated time for HITs can induce anxiety, according to one of the respondents.

*“I feel anxious when trying to race against the timer, even in non game type hits.”* P254 (Hearing impaired user)

**6.3.2 Perceived Depression.** According to the respondents’ feedback, a sense of depression could arise as a result of perceived difficulties with completing HITs and the difficulties associated with having to do crowdwork in addition to regular employment.

For instance, one of the respondents was apprehensive about selecting HITs because they were not sure that they could complete it in time.

*“Similar to the tasks I find difficult completing due to the amount of time, it makes me feel depressed that I’m more likely to return it and see the HITs appear as available and not bother even trying to accept it in the first place.”* P254 (Hearing impaired user)

P61 highlights pressures regarding working for most of the day and having little time for their personal life.

*“I have to wake up at 2:30 or 3AM to turk for an hour or two before I go work my actual job just to hopefully be able to pay all of my bills. And then I come home and turk.”* P61 (Cognitively impaired user.)

## 6.4 Relationships Between Disability Category and General Issues

In Table 10, we highlight the number of respondents in each disability category who reported issues in the general category, with respect to difficulties completing HITs due to time allocated, perceived anxiety and perceived depression.

### 6.5 Vision Issues

We asked the participants who reported having a visual impairment whether they found navigating the AMT interfaces difficult due to the visual layout; 54% (54/100) of visually-impaired respondents reported issues in this regard. One respondent’s comment provides insights into the factors responsible for these difficulties.

*“Some elements in surveys like selecting yes or no answer buttons or bullets seems so small to click. [...] We are in rush to finish the task within time. It is so difficult [to click those buttons] because we may miss clicking it. Those buttons must at least be average in size to [click them].”* P600

Although 87% of these respondents reported that they could complete an *Image Analysis* HIT with little difficulty, 58% reported difficulties completing other HITs as a result of their existing visual impairments.

When asked what could be done to alleviate these difficulties, one respondent suggested drawing on their social circle of fellow workers for support.

*“People are surprisingly willing to give support, if you ask for it in the right way. People performing similar tasks in other rooms could later exchange tips.”* P552

### 6.6 Hearing Issues

Most of the participants who reported hearing impairments did not have a serious condition, according to the results. For instance, 87% (87/100) of these respondents stated that they can hear loud speech in a quiet room; 77% stated that they could use a telephone without any special adaptations, and 73% reported that they could follow a conversation against background noise. Regarding completing HITs on AMT, 80% of respondents reported that they could complete an *Audio Transcription* HIT without difficulty. One of the respondents who found this HIT difficult to complete noted that their hearing impairment would not allow them to consider doing this HIT.

*“I don’t think anything can be done due to [the] type and degree of [my condition] (Tinnitus and Sensorineural hearing loss).”* P254

**Table 10: Respondents in each category, who report having other issues—difficulty completing HITs due to time, perceived anxiety and perceived depression.**

Category	Difficulty Completing HITs (%)	Perceived Anxiety (%)	Perceived Depression (%)
Cognition/Mental	47	45	44
Vision	59	56	54
Hearing	58	56	54
Motor/Dexterity	37	37	32
Reading	48	38	42

When asked whether the participants found any other HITs difficult to complete due to hearing impairments, 60% responded in the affirmative. Given that most of these respondents could complete an audio transcription task, it seems that issues with audio and hearing were more of a subtle nature. To this effect, P254 also elaborated:

*“[I find difficult, HITs] that require you to notice any other subtle changes in audio for videos...” P254*

### 6.7 Cognitive/Mental Issues

Approximately two thirds of participants who reported a cognitive disability appeared to have relatively mild conditions. For instance, 68% (68/100) of these respondents reported that they do not get confused about the time of day or who people are, and 79% could remember and pass on messages. However 45% reported that there were computer-based tasks that were difficult to perform as a result of cognitive issues, with 47% reporting difficulties completing HITs on AMT due to their existing cognitive impairment. The following comment highlights a respondent’s desire for additional time because they often get distracted while trying to complete HITs:

*“Some surveys don’t allow for enough time but I also would appreciate a longer time limit on most [HITs] because I have ADHD and I often get sidetracked.” P61*

### 6.8 Motor/Dexterity Issues

The findings were consistent regarding dexterity and motor abilities, as 77% (77/100) of these respondents reported being able to turn the knob of a cooker and pick up a pin without difficulty, and 69% stated that they could tie laces in a bow. Most of the participants in this category could plug in a USB cable and put on headphones (median = 4, interquartile range = 1) on the difficulty scale—1 = extreme difficulty; 5 = no difficulty.

Regarding using AMT, 77% of these respondents reported that they encountered difficulties when using AMT as a result of existing motor disabilities. No reasonable responses were given to the open-ended questions for this category; however, further insights can be gained in Section 7: Interview Findings.

### 6.9 Reading Issues

Two-fifths (40/100 or 40%) of the participants who reported having a reading disability stated that they encountered difficulties when completing HITs on AMT due to this disability. One of the respondents who reported having a reading disability highlights a lack

of sufficient instructions for a HIT, and he wishes the instructions could be simplified.

*“Some tasks must be simplified for the worker like adding detailed instructions on tasks to complete it properly.” P212*

## 7 INTERVIEW FINDINGS

Based on the emergent themes from our analysis, we discuss the findings from the interviews in this section. Following interviewee quotes, we suffix with their gender, age and reported disability category—for instance, female, 40, motor.

### 7.1 HITs Most Performed on AMT due to an Existing Disability

We asked the interviewees if there were any particular HITs that they performed mostly because of their disabilities. No particular HIT was mentioned as dominant in this case. However, all of the interviewees preferred completing surveys because they are relatively easy to do and offer better income given the effort. To this effect, P3 comments:

*“I try and do a survey because it takes less time and has a better time to money ratio.” P3 (male, 32, cognition)*

### 7.2 HITs Avoided on AMT due to an Existing Disability

We did not find a significant impact of disabilities on the types of HITs the interviewees most performed. However, since by nature, disabilities limit engagement, the interviewees were more opinionated regarding the type of HITs that were avoided as a result of their existing disability. For instance, P1 reports that he avoids HITs that contain sensitive content on depression and suicide ideation due to his own personal history.

*“[The HIT] might bring up things from my past that I don’t really want to get into, and it’s not really worth it to me to get into that emotional state for a dollar or fifty cents, so I’ll just drop them. A lot of times, there would be surveys with disturbing images, I would [normally] avoid those, but sometimes they don’t tell you that, and then you get into it and it’s like ‘well this isn’t really that great, I wish I didn’t see that one.’” P1 (male, 35, cognition)*

P2 avoids HITs that require a significant amount of time to complete due to a motor disability, as he comments:

*“I think HITs that are really long and they want you to pay attention for a very long time [...] sometimes my neck [goes into spasms]; so for people who have a problem like this because of their disability, it would be a nightmare for them.” P2 (male, 35, motor)*

P3 has difficulty understanding instructions, especially if they are numerous and non-simplified. For fear of having to return the HIT, on failure to complete it, he is apprehensive regarding accepting such HITs in the first place.

*“Sometimes if [a game type HIT] has so many instructions, it just makes me anxious because it is just like explaining it three different ways, I’m like ‘I’m gonna forget and then I’m gonna fail and then I’ll have to return the HIT.’” P3 (male, 32, cognition and hearing)*

P3 also suffers from Tinnitus; therefore, he finds it difficult to perform HITs that require processing audio.

*“I get tired really easily and that slows down my reading and comprehension speed. Because I have Tinnitus and Hearing loss, there’s a whole bunch of HITs like [audio transcription] or listening to a news article [as part of AI training] that can be difficult.” P3 (male, 32, cognition and hearing)*

P4 reflects on avoiding HITs that involve writing or typing, such as audio or video transcription, due to dyslexia.

*“I tend not to do much writing because I am terrible at writing; words can reverse themselves or numbers can just change position.” P4 (male, 38, cognition and reading)*

P4, who has severe difficulties in processing numbers, reflects on one occasion where, despite being a software developer, he returned a HIT that involved numerical data for training a conversational artificial intelligence (AI) agent.

*“I was trying the conversational AI HITs earlier, and it was where you have to follow up conversation and you’re supposed to use the knowledge and rearrange it. It got to a point [where there were] a lot of numbers; that’s where I quit.” P4 (male, 38, cognition and reading)*

### 7.3 Qualifications

There was a general uncertainty around the requirements for a masters qualification on AMT. None of the interviewees had achieved the qualification despite using the AMT platform for at least a year, and none of them knew how to obtain the masters qualification. This was an issue since the qualification allows workers to gain access to higher paid HITs. P5’s and P4’s comments echo the general sentiment regarding this problem.

*“[AMT] does not hand out their masters qualification, they just don’t pay attention to it. The requesters come in and say ‘I want the best workers’, and no one has [the masters].” P5 (female, 35, cognition)*

*“No one knows how to get qualifications such as masters.” P4 (male, 38, cognition and reading)*

### 7.4 Time to Complete HITs

The survey responses highlighted the fact that the allocated time for HITs can present a significant barrier for the participants. In the interview, P2 responds that the allocated time often presents a barrier to HIT completion mostly due to his impairment. Given that worker reputation can be affected by unfinished HITs, this can present a significant problem for AMT workers.

*“My jaw pops in and out, so if it’s dislocated, my neck will start pulling, and it makes it very hard to look at the computer, so I need to take a break. If I have a lot of HITs in my queue and I’m on a timer, and I’m in an instance where my symptoms flare up, I might have to*

*return some of them, because I just can’t do them in the time allotted.” P2 (male, 35, motor)*

### 7.5 Wages and Income

All of the interviewees stated that they engage in crowdwork on AMT as a means to earn supplemental income, since they are currently permanently employed. However, none of the interviewees said that the wages were enough given the HITs that they normally performed. P5, who reported suffering from severe anxiety, highlighted experiencing stress in certain situations where she feels unfairly treated through HIT rejections.

*“A lot of people like to just take your information and go. You put in all this work and you get rejected, so not only do you not get paid but it is a huge negative on your account and it signals other [requesters] that you are a crappy worker, because this one person decided to reject you. It’s really stressful because I could use the extra money [in my current circumstances].” P5 (female, 35, cognition)*

### 7.6 Goals

Three of the interviewees stated that they set goals when using AMT, with *income* being the key focus of the goal. For P5, who suffers from post-traumatic stress disorder (PTSD) and attention deficit hyperactivity disorder (ADHD), work goals had a more serious effect on anxiety.

*“I am just a giant ball of anxiety. On a scale of one to ten, I’m always really a ten. It’s so frustrating because I know from my personal budget that if I don’t hit a goal of at least \$10 every day, I definitely won’t have enough money [for daily life].” P5 (female, 35, cognition)*

### 7.7 User Interface (UI)

When asked about user interface problems, the interviewees felt that as experienced AMT workers, the user interface did not present a problem. However, they felt that they struggled to use the interface as less experienced workers and that new workers would also face the same problems. For instance, P1 highlights issues with the user interface.

*“[AMT] has a bit of a learning curve to it in order to actually be productive in a way that is in any way shape or form meaningful. Somebody who knows nothing about it, goes to the [AMT] page registers as a worker and then starts trying to work HITs from the [AMT] user interface is going to have a really bad time, because it’s hard to find HITs that pay well.” P1 (male, 35, cognition)*

P1 also imagines what certain UI elements, such as a CAPTCHA could present for people with a reading disability.

*“[There was] a very extreme CAPTCHA...it had like letters in front of letters [...] If someone had reading problems that would have been a little bit hard.” P1 (male, 35, cognition)*

The interviewees were asked to comment on potential usability issues using a screenshot of the AMT user homepage, which displays available HITs. Most responses indicated that the interviewees were either used to the platform or used scripts to access HITs (see Section 7.8.3 below), therefore they barely encountered the user home page. However, P3 highlights potential difficulties regarding comprehending instructions; P3 suggests redesigning the interface to allow users to hide and view certain interface elements.

*“I think there’s a little too much information on the screen at the same time. I think that it should be optional what you want to see.”*

**P3 (male, 32, cognition and hearing)**

## 7.8 Improving Accessibility in AMT

We asked the interviewees to comment on what Amazon and third party sites owned by requesters could do to alleviate the key accessibility issues that they face as AMT workers. We categorize and discuss their responses in this section.

**7.8.1 Standardized Platforms and Tools.** P1 felt that there should be a standard set of guidelines and ways to outline certain HITs that are similar, for instance surveys.

*“At least for basic surveys, it would be nice if they were hosted on the same general platform, because that will help with efficiency and also with confusion sometimes.”* **P1 (male, 35, cognition)**

**7.8.2 Disability-Enabled Profiles.** In response to encouraging a fairer system, P1 and P2 suggest using profiles to distinguish between workers with and without disabilities. For those workers who have disabilities, accommodations can be made to enable them to complete HITs with less difficulty. P1 draws comparisons between AMT and Prolific (a UK-based crowdwork platform) [50], which allows for greater flexibility in tracking HIT time.

*“If you could update your profile to include that you have disabilities and it can somehow flag your account and to let requesters know and maybe a requester can select an option like ‘this person needs a little bit more time’, and for people who do have disabilities, their account will be flagged as such. So they would see the extended time and the [non-disabled] people wouldn’t, that’s something that I think could help.”* **P2 (male, 35, motor)**

*“If there was a standard way within the platform to say this is the estimated time that this should take, and that the platform enforced that you should at least give some kind of multiplier to the estimated time [...] On Prolific there’s kind of a difference between when you accept the task and when you start working on the task, whereas on AMT, you accept the task and then the timer starts... I don’t know if Amazon cares to make any changes to it.”* **P1 (male, 35, cognition)**

**7.8.3 Tools.** We already highlighted, in our study findings, user frustration regarding finding HITs that offer a reasonable income reward. In response to these problems, several tools, such as TurkerView, HIT Catcher and MTurk Suite, have emerged, which allow workers to manage their AMT account and find reasonable HITs. Based on interviewee feedback these tools have become an essential part of their crowdwork workflow.

*“TurkerView could give you the requester name, the HIT title, it fills in the wage data, how long it took me to complete it; it has five ratings where I could go from underpaid to generous, and it lets me fill in how much per hour that HIT got me [...] I also used a tool called Cubicle and it’s beautiful. It’s practically earned me triple what I made before. I have to pay \$20 a month or it, but it allows me to get to where I want to be [financially].”* **P5 (female, 35, cognition)**

*“MTurk Suite is great, and when I found that, I said ‘this is awesome, it will auto search for things and I will tell it any time this HIT from a particular requester comes in, grab it for me’, so I will spend more time working and less time searching.”* **P1 (male, 35, cognition)**

*“Maybe if [Amazon] can build their own HIT Finder and HIT Catcher, that will be the best thing that they can do really. These [tools] are really big quality of life improvements, because you don’t really get them with the regular [AMT] dashboard.”* **P2 (male, 35, motor)**

**7.8.4 Community Support.** In addition to tools, as highlighted above, community support is also essential to gaining an advantage as a crowdworker on AMT. For instance, P2 comments on the importance of communities for learning about tools that could give them advantage in crowdwork.

*“If you join communities and take some time to learn the platform with these tools, it could be very, very lucrative. Luckily I was able to get in these communities and read resources to where I could understand, but Amazon does not really do enough to tell you how you can make AMT worth your while [...] It’s not intuitive.”* **P2 (male, 35, motor)**

Given, from our findings, that AMT can be perceived as difficult to use (for new workers) and it requires an effort be made to search for crowdwork communities, we asked the interviewees what the community currently does to reach out to new workers who may need help. P2 responds with the following reflection:

*“Yes and also no, [the community does not reach out to new workers]. For me, I like to look up stuff [...], but for someone who’s not like that it’s going to be very difficult for them to find these resources.”* **P2 (male, 35, motor)**

All but one of the interviewees were active members of one or more crowdwork communities. P4, being a software developer by profession, highlights the importance of participating in the community by helping fellow workers with software tools to enable effective better paid crowdwork.

*“I use Turkernator, which is created by a fellow Turker; he also created MTurk Suite. I worked closely with him in developing Turkernator. I basically have all of my settings to where I only see the cream of the crop type HITs. I can just click on a HIT and be directed to TurkerView to see whether it is worth my time or not.”* **P4 (male, 38, cognition and reading)**

## 8 DISCUSSION

We discuss our study findings in light of our research objectives—highlighted in Section 3.

### 8.1 Demographics of AMT Workers with Disabilities

Through our first objective, we sought to highlight the proportion of Mechanical Turk (AMT) workers that identify as having a disability. Our findings are consistent with the literature in the sense that the average worker is relatively young, educated and employed [17, 25, 34, 49]. However, in our study there seems to be a higher representation of males (57.8%), compared to the higher representation of females indicated in the literature (70% in 2009 [34] and 50+% more recently in 2018 [17, 26]), although this could be indicative of the fact that more males responded to the topic of our research in the first survey HIT.

Our findings are also consistent with the literature [26] as we found that, although income was an important reason for participating in crowdwork, AMT workers also do crowdwork for a variety

of other reasons such as keeping productive, performing enjoyable tasks and contributing to research. Income was not seen as sufficient for most HITs [25, 26] with most survey respondents using crowdwork as a supplemental source of income. However, our participants use tools, such as MTurk Suite, HIT Finder and Turkernator to identify high income HITs and manage their crowdwork workflow. The key benefit of crowdwork, according to our findings, is that work can be undertaken from home, enabling a flexible and potentially less stressful form of work [60, 61].

## 8.2 Accessibility Issues in AMT

In our second research objective, we sought to identify how AMT workers are limited by their individual disabilities and how accessibility issues may be exacerbated given specific disabilities, which are likely to impact computer use—vision, cognition/mental, hearing, reading and motor/dexterity [16]. Two thirds of the respondents from our first survey reported having at least one disability, or difficulty using a computer, thereby presenting potential barriers to completing crowdwork tasks. This has significant implications for HCI, and it is not possible to exhaust all of the issues in a single piece of work. Nevertheless, we highlight the key issues as they relate to the individual disability categories.

Our work echoes the findings from previous work [61] that the allocated time for crowdwork tasks (or HITs) can be a significant barrier for crowdworkers, resulting in increased anxiety and a feeling of depression if HITs cannot be completed in time. In numerous cases, the need for additional time is essential given the limitations afforded by disabilities, such as motor/dexterity or cognition/mental. Key contributors to anxiety include the repercussions of failing to complete HITs on worker reputation, which affects future opportunities for work on the platform.

A relatively high number of respondents in Survey 2 (37% to 77%) indicated that they encountered some sort of difficulty when completing HITs on AMT, due to their disability. Workers with motor/dexterity impairments reported the highest occurrence in this regard (77%), which supports the study by Lund et al. [43] that found physical disabilities to be the most prevalent category. We note in the findings that most of the issues are also applicable to the general population, for example, wages, HIT completion time, ambiguity regarding qualifications and the user interface. However, the fact that these issues can be significantly exacerbated due to the presence of disabilities, as we report through our findings, is noteworthy, and this is an important starting point towards improving the design of crowdwork platforms.

## 8.3 Implications for Design

In our third research objective, we sought to discuss implications for the design of crowdwork platforms and tasks, to enable effective use by people who have disabilities. We found numerous accessibility issues, in both Survey 2 and the interviews, and we consequently discuss implications for design in this section.

**8.3.1 Streamlining Workflows in Crowdwork.** We highlight difficulties regarding engaging in crowdwork arising from the following: a) a lack of information, for new and experienced workers, on how to find HITs that are lucrative for the worker; b) rigidity in allocated HIT time, especially in certain HITs posted by relatively new

requesters, which presents issues for workers with disabilities; and c) a lack of clarity on how to achieve qualifications necessary to access HITs with higher payments. We suggest that crowdwork platform holders, such as Amazon, streamline the workflow for microtasks, to reduce worker inequities in crowdwork.

To mitigate the above issues, we first suggest that crowdwork platform holders (e.g. Amazon) ensure that instructions are clear, to reduce confusion and promote usability. To mitigate the risk of overloading the user interface, information can be contextualized based on user preferences and relevance. Further, we suggest that suitable warnings are issued on AMT in cases where sensitive material exists—for instance, HITs that discuss suicide ideation (for workers who have a history in this regard) and HITs that involve heavy numerical data (for workers who suffer from dyscalculia).

Second, Zyskowski et al. [61] suggest platform holders use meta-data to indicate the kinds of abilities that are suitable for HITs; this could work in the case of popular HITs for which it is obvious that certain abilities are essential, for example, having no hearing impairment when doing an audio transcription HIT. However, it is possible that this solution could lead to complications in more nuanced cases, as found in our work. For instance, a requester may need to specify all of the relevant abilities for their HITs, and there are numerous conditions to cater to, some of which the requester may not even be aware of, such as dyscalculia in AI training tasks. Based on our user feedback, we recommend an additional solution where platform holders take responsibility and integrate such flexibility into user profiles. For instance, if a worker has a certain disability, as stated in their profile, they do not get penalized if they cannot complete HITs that might be difficult for them in the allocated time. Instead they are given extra time to complete HITs. As a general rule, the approach taken by Prolific [50] can be considered in AMT—activating the HIT timer only when a worker begins working on the HIT, rather than when the HIT is accepted.

Third, in the case of AMT, Amazon should also clarify for workers how to achieve qualifications, such as the *masters*, since this is necessary to access higher-paying HITs. One way to do this is to provide some indication of how much progress a worker has made towards achieving certain qualifications. This could potentially encourage workers to perform more HITs in order to become more qualified on the platform.

**8.3.2 Implementing Useful Functionality.** Our survey respondents heavily relied on external tools and scripts to assist them in managing their HITs, for example, finding and tracking reasonable HITs from reputable requesters. We suggest that Amazon implements important functions provided by these third party tools into the AMT platform in order to promote usability. New workers and workers with disabilities will benefit greatly from this suggestion.

**8.3.3 User Interface Redesign.** AMT workers in our study remarked that the AMT interface does not regularly get updated, hence they resort to external tools to manage HITs. Using a screenshot of the worker home page, the interviewees identified potential barriers to usability, such as presenting too much information and using relatively small interface elements, for instance buttons and hyperlinks. Calvo et al. [10], in their 2014 heuristic evaluation of the AMT interface, noted several heuristic violations which could have

implications for usability. We suggest that platform holders conduct such evaluations in addition to user studies to reduce potential barriers to effective use.

**8.3.4 Leveraging the Community.** According to our findings, communities offer essential support to AMT workers by highlighting useful information on available tools, tips, HITs and social benefits [32]. It can be difficult for new workers to access such communities, mostly due to a lack of awareness. On the one hand, platform holders can promote communities, such as Turker Nation, especially to new workers on the platform. On the other hand communities can, perhaps, make more of an effort to reach out to new workers or workers with disabilities, to help them on their crowdwork journey.

## 9 LIMITATIONS

First, our work focused on US workers. Given the increasing prevalence of international workers on the AMT platform, especially in India, opportunities exist for further work to explore accessibility in crowdwork with participants from other countries, which we expect could differ from our findings, given cultural differences.

Second, five participants responded to our interview call. Our interviewees responded in great depth to accessibility issues that they often face as AMT workers, with regard to their disabilities; this provides useful insights into improving accessibility in crowdwork platforms. However, their opinions represent a subset of the potential barriers to usability in crowdwork platforms in the domain of the explored disability categories. We note here that we experienced difficulties recruiting participants for the interview. Similar problems are evident in the literature. For example, Sannon and Cosley [54] interviewed 14 regular AMT participants over a 4-month period. The exact factors contributing to low interview uptake are unclear. We conjecture factors such as anonymity and convenience to play a role, based on anecdotal evidence. For instance, crowdworkers can complete HITs in their own time and pace, and the work is relatively anonymous. On the other hand, an interview requires some degree of organization, rigidity and exposure.

Third, due to the focus of our work, we explored a set of disability dimensions which are likely to impact computer use [16]. However, there are potentially other disability categories that may negatively impact usability in crowdwork tasks and can be explored in this area, for example, the autism spectrum. This is an opportunity for further work.

## 10 CONCLUSIONS

Understanding the attitudes and work preferences of crowdworkers is a valuable necessary initial step towards improving design [24]. We sought to identify how crowdworkers with disabilities cope with performing Human Intelligence Tasks (HITs) on the Amazon Mechanical Turk (AMT) platform. We conducted two surveys to understand the key challenges for completing HITs, which could present barriers for effective usability. Following up with interviews, we found that accessibility issues pertaining to various disabilities can affect which HITs crowdworkers perform and avoid.

Based on our findings, we highlighted implications for design regarding the AMT platform, many of which could extend to other crowdwork platforms, in order to promote usability and reduce

user frustration, especially for workers with disabilities. The most important solutions in this regard include: streamlining crowdwork workflows, implementing functionality from third party tools, a user interface redesign and leveraging community benefits.

Our work constitutes one important step towards improving accessibility in crowdwork, which is an increasingly important avenue for work, given its potential benefits for users with disabilities, such as work flexibility, additional income, avoiding work commutes and increasing work-from-home opportunities—a necessary option during the current Covid-19 pandemic.

## ACKNOWLEDGMENTS

This work was supported by the EPSRC (grants EP/R004471/1 and EP/S027432/1). Supporting data for this publication is available at <https://doi.org/10.17863/CAM.62937>.

## REFERENCES

- [1] Harini Alagarai Sampath, Rajeev Rajeshuni, and Bipin Indurkha. 2014. Cognitively inspired task design to improve user performance on crowdsourcing platforms. In *Proceedings of the 32nd annual ACM conference on Human factors in computing systems - CHI '14*. ACM Press, Toronto, Ontario, Canada, 3665–3674. <https://doi.org/10.1145/2556288.2557155>
- [2] Amazon. 2020 (accessed September 11, 2020). Amazon Mechanical Turk. <https://www.mturk.com/>
- [3] Daniel Archambault, Helen C. Purchase, and Tobias Hobfeld. 2017. Evaluation in the Crowd: An Introduction. In *Evaluation in the Crowd. Crowdsourcing and Human-Centered Experiments*, Daniel Archambault, Helen Purchase, and Tobias Hobfeld (Eds.). Vol. 10264. Springer International Publishing, Cham, 1–5. [https://doi.org/10.1007/978-3-319-66435-4\\_1](https://doi.org/10.1007/978-3-319-66435-4_1)
- [4] Giorgio Brajnik, Yeliz Yesilada, and Simon Harper. 2010. Testability and validity of WCAG 2.0: the expertise effect. In *Proceedings of the 12th international ACM SIGACCESS conference on Computers and accessibility - ASSETS '10*. ACM Press, Orlando, Florida, USA, 43. <https://doi.org/10.1145/1878803.1878813>
- [5] Giorgio Brajnik, Yeliz Yesilada, and Simon Harper. 2012. Is accessibility conformance an elusive property? A study of validity and reliability of WCAG 2.0. *ACM Transactions on Accessible Computing* 4, 2 (March 2012), 1–28. <https://doi.org/10.1145/2141943.2141946>
- [6] Matthew W. Brault. 2012. Americans With Disabilities: 2010. <https://www2.census.gov/library/publications/2012/demo/p70-131.pdf>
- [7] Alice M. Brawley and Cynthia L.S. Pury. 2016. Work experiences on MTurk: Job satisfaction, turnover, and information sharing. *Computers in Human Behavior* 54 (Jan. 2016), 531–546. <https://doi.org/10.1016/j.chb.2015.08.031>
- [8] Robin Brewer, Meredith Ringel Morris, and Anne Marie Piper. 2016. "Why would anybody do this?": Understanding Older Adults' Motivations and Challenges in Crowd Work. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*. ACM, San Jose California USA, 2246–2257. <https://doi.org/10.1145/2858036.2858198>
- [9] US Census Bureau. 2020. U.S. Census Bureau QuickFacts: United States. <https://www.census.gov/quickfacts/fact/table/US/PST045219>
- [10] Rocío Calvo, Shaun K. Kane, and Amy Hurst. 2014. Evaluating the accessibility of crowdsourcing tasks on Amazon's mechanical turk. In *Proceedings of the 16th international ACM SIGACCESS conference on Computers & accessibility - ASSETS '14*. ACM Press, Rochester, New York, USA, 257–258. <https://doi.org/10.1145/2661334.2661401>
- [11] Krista Casler, Lydia Bickel, and Elizabeth Hackett. 2013. Separate but equal? A comparison of participants and data gathered via Amazon's MTurk, social media, and face-to-face behavioral testing. *Computers in Human Behavior* 29, 6 (Nov. 2013), 2156–2160. <https://doi.org/10.1016/j.chb.2013.05.009>
- [12] Jesse Chandler and Danielle Shapiro. 2016. Conducting Clinical Research Using Crowdsourced Convenience Samples. *Annual Review of Clinical Psychology* 12, 1 (March 2016), 53–81. <https://doi.org/10.1146/annurev-clinpsy-021815-093623>
- [13] Yanto Chandra and Liang Shang. 2019. Inductive Coding. In *Qualitative Research Using R: A Systematic Approach*. Springer Singapore, Singapore, 91–106. [https://doi.org/10.1007/978-981-13-3170-1\\_8](https://doi.org/10.1007/978-981-13-3170-1_8)
- [14] Cint. 2020 (accessed September 11, 2020). The Cint Platform. <https://www.cint.com/platform-market-research-technology>
- [15] Clickworker. 2020 (accessed September 11, 2020). Clickworker Website. <https://www.clickworker.com/>
- [16] Microsoft Corporation. 2003. *The Wide Range of Abilities and Its Impact on Computer Technology*. Technical Report. Forrester Research Inc. <https://www.microsoft.com/en-us/download/details.aspx?id=18446>

- [17] 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 - WSDM '18*. ACM Press, Marina Del Rey, CA, USA, 135–143. <https://doi.org/10.1145/3159652.3159661>
- [18] Jennifer Fereday and Eimear Muir-Cochrane. 2006. Demonstrating Rigor Using Thematic Analysis: A Hybrid Approach of Inductive and Deductive Coding and Theme Development. *International Journal of Qualitative Methods* 5, 1 (March 2006), 80–92. <https://doi.org/10.1177/160940690600500107>
- [19] Avi Fleischer, Alan D. Mead, and Jialin Huang. 2015. Inattentive Responding in MTurk and Other Online Samples. *Industrial and Organizational Psychology* 8, 2 (June 2015), 196–202. <https://doi.org/10.1017/iop.2015.25>
- [20] D. Jake Follmer, Rayne A. Sperling, and Hoi K. Suen. 2017. The Role of MTurk in Education Research: Advantages, Issues, and Future Directions. *Educational Researcher* 46, 6 (Aug. 2017), 329–334. <https://doi.org/10.3102/0013189X17725519>
- [21] Karèn Fort, Gilles Adda, and K. Bretonnel Cohen. 2011. Amazon Mechanical Turk: Gold Mine or Coal Mine? *Computational Linguistics* 37, 2 (June 2011), 413–420. [https://doi.org/10.1162/COLL\\_a\\_00057](https://doi.org/10.1162/COLL_a_00057)
- [22] 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 - HT '14*. ACM Press, Santiago, Chile, 218–223. <https://doi.org/10.1145/2631775.2631819>
- [23] Ujwal Gadiraju, Sebastian Müller, Martin Nöllenburg, Dietmar Saupe, Sebastian Egger-Lampf, Daniel Archambault, and Brian Fisher. 2017. Crowdsourcing Versus the Laboratory: Towards Human-Centered Experiments Using the Crowd. In *Evaluation in the Crowd. Crowdsourcing and Human-Centered Experiments*, Daniel Archambault, Helen Purchase, and Tobias Hobfeld (Eds.). Vol. 10264. Springer International Publishing, Cham, 6–26. [https://doi.org/10.1007/978-3-319-66435-4\\_2](https://doi.org/10.1007/978-3-319-66435-4_2)
- [24] Neha Gupta, David Martin, Benjamin V. Hanrahan, and Jacki O'Neill. 2014. TurkLife in India. In *Proceedings of the 18th International Conference on Supporting Group Work - GROUP '14*. ACM Press, Sanibel Island, Florida, USA, 1–11. <https://doi.org/10.1145/2660398.2660403>
- [25] 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 - CHI '18*. ACM Press, Montreal QC, Canada, 1–14. <https://doi.org/10.1145/3173574.3174023>
- [26] Kotaro Hara, Abigail Adams, Kristy Milland, Saiph Savage, Benjamin V. Hanrahan, Jeffrey P. Bigham, and Chris Callison-Burch. 2019. Worker Demographics and Earnings on Amazon Mechanical Turk: An Exploratory Analysis. In *Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems*. ACM, Glasgow Scotland UK, 1–6. <https://doi.org/10.1145/3290607.3312970>
- [27] Kotaro Hara, Vicki Le, and Jon Froehlich. 2013. Combining crowdsourcing and google street view to identify street-level accessibility problems. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems - CHI '13*. ACM Press, Paris, France, 631. <https://doi.org/10.1145/2470654.2470744>
- [28] David J. Hauser and Norbert Schwarz. 2016. Attentive Turkers: MTurk participants perform better on online attention checks than do subject pool participants. *Behavior Research Methods* 48, 1 (March 2016), 400–407. <https://doi.org/10.3758/s13428-015-0578-z>
- [29] Isabell Hensel, Jochen Koch, and Eva Kocher. 2016. Crowdsourcing als Phänomen der Koordination digitaler Erwerbsarbeit - Eine interdisziplinäre Perspektive. *Industrielle Beziehungen* 2 (2016), 162–186. <https://doi.org/10.1688/IndB-2016-02-Hensel>
- [30] Matthias Hirth, Jason Jacques, Peter Rodgers, Ognjen Scekcic, and Michael Wybrow. 2017. Crowdsourcing Technology to Support Academic Research. In *Evaluation in the Crowd. Crowdsourcing and Human-Centered Experiments*, Daniel Archambault, Helen Purchase, and Tobias Hobfeld (Eds.). Vol. 10264. Springer International Publishing, Cham, 70–95. [https://doi.org/10.1007/978-3-319-66435-4\\_4](https://doi.org/10.1007/978-3-319-66435-4_4)
- [31] Paul Hitlin. 2016. *Research in the Crowdsourcing Age, a Case Study: How Scholars, Companies and Workers are Using Mechanical Turk, a "gig Economy" Platform, for Tasks Computers Can't Handle*. Pew Research Center.
- [32] Hwajung Hong, Svetlana Yarosh, Jennifer G. Kim, Gregory D. Abowd, and Rosa I. Arriaga. 2013. Investigating the use of circles in social networks to support independence of individuals with autism. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems - CHI '13*. ACM Press, Paris, France, 3207. <https://doi.org/10.1145/2470654.2466439>
- [33] J Howe. 2006. The Rise of Crowdsourcing. *Wired Magazine* 14, 6 (2006). <https://www.wired.com/2006/06/crowds/>
- [34] Panos Ipeirotis. 2009 (accessed September 11, 2020). *Turker Demographics vs Internet Demographics*. <https://www.behind-the-enemy-lines.com/2009/03/turker-demographics-vs-internet.html>
- [35] Jason T. Jacques and Per Ola Kristensson. 2017. Design Strategies for Efficient Access to Mobile Device Users via Amazon Mechanical Turk. In *Proceedings of the First ACM Workshop on Mobile Crowdsensing Systems and Applications - CrowdSensSys '17*. ACM Press, Delft, Netherlands, 25–30. <https://doi.org/10.1145/3139243.3139247>
- [36] Jason T. Jacques and Per Ola Kristensson. 2019. Crowdsourcing Economics in the Gig Economy. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems - CHI '19*. ACM Press, Glasgow, Scotland UK, 1–10. <https://doi.org/10.1145/3290605.3300621>
- [37] Shashank Khanna, Aishwarya Ratan, James Davis, and William Thies. 2010. Evaluating and improving the usability of Mechanical Turk for low-income workers in India. In *Proceedings of the First ACM Symposium on Computing for Development - ACM DEV '10*. ACM Press, London, United Kingdom, 1. <https://doi.org/10.1145/1926180.1926195>
- [38] Aniket Kittur, Ed H. Chi, and Bongwon Suh. 2008. Crowdsourcing user studies with Mechanical Turk. In *Proceeding of the twenty-sixth annual CHI conference on Human factors in computing systems - CHI '08*. ACM Press, Florence, Italy, 453. <https://doi.org/10.1145/1357054.1357127>
- [39] Masamoto Kobayashi, Tatsuya Ishihara, Akihiro Kosugi, Hironobu Takagi, and Chieko Asakawa. 2013. Question-Answer Cards for an Inclusive Micro-tasking Framework for the Elderly. In *Human-Computer Interaction - INTERACT 2013*, David Hutchison, Takeo Kanade, Josef Kittler, Jon M. Kleinberg, Friedemann Mattern, John C. Mitchell, Moni Naor, Oscar Nierstrasz, C. Pandu Rangan, Bernhard Steffen, Madhu Sudan, Demetri Terzopoulos, Doug Tygar, Moshe Y. Vardi, Gerhard Weikum, Paula Kotze, Gary Marsden, Gittie Lindgaard, Janet Wesson, and Marco Winckler (Eds.). Vol. 8119. Springer Berlin Heidelberg, Berlin, Heidelberg, 590–607. [https://doi.org/10.1007/978-3-642-40477-1\\_38](https://doi.org/10.1007/978-3-642-40477-1_38)
- [40] Siou Chew Kuek and Cecilia Paradi-Guildford. 2015. *The Global Opportunity in Online Outsourcing*. Technical Report. World Bank Group. <http://documents1.worldbank.org/curated/en/13837146800900555/pdf/ACS14228-ESW-white-cover-P149016-Box391478B-PUBLIC-World-Bank-Global-OO-Study-WB-Rpt-Finals.pdf>
- [41] Leib Litman, Jonathan Robinson, and Cheskie Rosenzweig. 2015. The relationship between motivation, monetary compensation, and data quality among US- and India-based workers on Mechanical Turk. *Behavior Research Methods* 47, 2 (June 2015), 519–528. <https://doi.org/10.3758/s13428-014-0483-x>
- [42] Eric Loepp and Jarrod T. Kelly. 2020. Distinction without a difference? An assessment of MTurk Worker types. *Research & Politics* 7, 1 (Jan. 2020), 205316801990118. <https://doi.org/10.1177/2053168019901185>
- [43] Emily M. Lund, Michael R. Nadorff, Kate Galbraith, and Katie B. Thomas. 2018. Using Amazon Mechanical Turk to Recruit Participants With Disabilities. *SAGE Research Methods Cases in Psychology* (2018). <https://doi.org/10.4135/9781526437280>
- [44] Adam Marcus and Aditya Parameswaran. 2015. Crowdsourced Data Management: Industry and Academic Perspectives. *Foundations and Trends in Databases* 6, 1-2 (2015), 1–161. <https://doi.org/10.1561/19000000044>
- [45] 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 - CSCW '14*. ACM Press, Baltimore, Maryland, USA, 224–235. <https://doi.org/10.1145/2531602.2531663>
- [46] Winter Mason and Siddharth Suri. 2012. Conducting behavioral research on Amazon's Mechanical Turk. *Behavior Research Methods* 44, 1 (March 2012), 1–23. <https://doi.org/10.3758/s13428-011-0124-6>
- [47] A. Moghaddam. 2006. Coding issues in grounded theory. *Issues in Educational Research* 16 (01 2006), 47–58.
- [48] Gabriele Paolacci and Jesse Chandler. 2014. Inside the Turk: Understanding Mechanical Turk as a Participant Pool. *Current Directions in Psychological Science* 23, 3 (June 2014), 184–188. <https://doi.org/10.1177/0963721414531598>
- [49] Gabriele Paolacci, Jesse Chandler, and Panagiotis G. Ipeirotis. 2010. Running experiments on Amazon Mechanical Turk. *Decision Making* 5, 5 (2010), 411–419.
- [50] Prolific. 2020 (accessed September 11, 2020). Prolific. <https://www.prolific.co/>
- [51] Elissa M. Redmiles, Sean Kross, and Michelle L. Mazurek. 2019. How Well Do My Results Generalize? Comparing Security and Privacy Survey Results from MTurk, Web, and Telephone Samples. In *2019 IEEE Symposium on Security and Privacy (SP)*. IEEE, San Francisco, CA, USA, 1326–1343. <https://doi.org/10.1109/SP.2019.00014>
- [52] Mireia Ribera, Merce Porras, Marc Boldu, Miquel Termens, Andreu Sule, and Pilar Paris. 2009. Web Content Accessibility Guidelines 2.0: A further step towards accessible digital information. *Program* 43, 4 (Sept. 2009), 392–406. <https://doi.org/10.1108/00330330910998048>
- [53] Joel Ross, Lilly Irani, M. Six Silberman, Andrew Zaldivar, and Bill Tomlinson. 2010. Who are the crowdworkers?: shifting demographics in mechanical turk. In *Proceedings of the 28th of the international conference extended abstracts on Human factors in computing systems - CHI EA '10*. ACM Press, Atlanta, Georgia, USA, 2863. <https://doi.org/10.1145/1753846.1753873>
- [54] Shruti Sannon and Dan Cosley. 2019. Privacy, Power, and Invisible Labor on Amazon Mechanical Turk. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems - CHI '19*. ACM Press, Glasgow, Scotland UK, 1–12. <https://doi.org/10.1145/3290605.3300512>
- [55] Kathryn Sharpe Wessling, Joel Huber, and Oded Netzer. 2017. MTurk Character Misrepresentation: Assessment and Solutions. *Journal of Consumer Research* 44, 1 (June 2017), 211–230. <https://doi.org/10.1093/jcr/ucx053>
- [56] Scott M. Smith, Catherine A. Roster, Linda L. Golden, and Gerald S. Albaum. 2016. A multi-group analysis of online survey respondent data quality: Comparing a regular USA consumer panel to MTurk samples. *Journal of Business Research* 69,

- 8 (Aug. 2016), 3139–3148. <https://doi.org/10.1016/j.jbusres.2015.12.002>
- [57] Rion Snow, Daniel Jurafsky, and Andrew Y. Ng. 2020 (accessed September 11, 2020). Cheap and Fast - But is it Good? Evaluating Non-Expert Annotations for Natural Language Tasks.
- [58] Alexander Sorokin and David Forsyth. 2008. Utility data annotation with Amazon Mechanical Turk. In *2008 IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops*.
- [59] Neil Stewart, Christoph Ungemach, Adam J. L. Harris, Daniel M. Bartels, Ben R. Newell, Gabriele Paolaccik, and Jesse Chandler. 2015. The average laboratory samples a population of 7,300 Amazon Mechanical Turk workers. *Judgment and Decision Making* 10, 5 (2015), 479–491.
- [60] Saiganesh Swaminathan, Kotaro Hara, and Jeffrey P. Bigham. 2017. The Crowd Work Accessibility Problem. In *Proceedings of the 14th Web for All Conference on The Future of Accessible Work - W4A '17*. ACM Press, Perth, Western Australia, Australia, 1–4. <https://doi.org/10.1145/3058555.3058569>
- [61] Kathryn Zyskowski, Meredith Ringel Morris, Jeffrey P. Bigham, Mary L. Gray, and Shaun K. Kane. 2015. Accessible Crowdwork?: Understanding the Value in and Challenge of Microtask Employment for People with Disabilities. In *Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing - CSCW '15*. ACM Press, Vancouver, BC, Canada, 1682–1693. <https://doi.org/10.1145/2675133.2675158>