Introducing People with ASD to Crowd Work

Kotaro Hara
HCI and Robotics Institutes
Carnegie Mellon University
kotaroh@cmu.edu

Jeffrey P. Bigham
HCI and LT Institutes
Carnegie Mellon University
jbigham@cs.cmu.edu

ABSTRACT
Adults with Autism Spectrum Disorders (ASD) are unemployed at a high rate, in part because the constraints and expectations of traditional employment can be difficult for them. In this paper, we report on our work in introducing people with ASD to remote work on a crowdsourcing platform and a prototype tool we developed by working with participants. We conducted a six-week long user-centered design study with three participants with ASD. The early stage of the study focused on assessing the abilities of our participants to search and work on micro-tasks available on the crowdsourcing market. Based on our preliminary findings, we designed, developed, and evaluated a prototype tool to facilitate image transcription tasks that are increasingly popular on crowd labor markets. Our findings suggest that people with ASD have varying levels of ability to work on micro-tasks, but are likely to be able to work on tasks like image transcription. The tool we introduce, Assistive Task Queue (ATQ), facilitated our participants' completion of image transcription tasks by removing ambiguity in finding the next task to work on and in simplifying tasks into discrete steps. ATQ may serve as a general platform for finding and delivering appropriate tasks to workers with autism.

CCS Concepts
• Human-centered computing → Accessibility →Accessibility technologies.

Keywords
Autism Spectrum Disorder; crowdsourcing

1. INTRODUCTION
Autism Spectrum Disorder (ASD) is a developmental disorder that is characterized by communication deficiencies, social interaction challenges, and restricted, repetitive patterns of behavior (*e.g.,* stimming) with varying levels of severity [36, 39]. In the U.S., more than 3.5 million people are estimated to have ASD [1], of which 50-75% of adults are expected to be unemployed [15]—a rate much higher than people without disabilities [6]. The lack of job opportunities not only negatively impacts the independence and self-esteem of the affected population, but also has a broader economic effect [8, 23].

Prior work in ASD has largely focused on understanding the causes of autism [35], designing intervention methods and technologies to reduce the impact of disorders on people with ASD [24], and creating technologies to support children and their caregivers [13, 21]. However, we know little about the challenges that adults with ASD in traditional workplace face, and we know even less on how technology can mitigate the vocational challenges for individuals with ASD. Recent work has examined how people with ASD work, but this was focused on high-functioning individuals who work in the technology-industry [30]. Zyskowski *et al.* [42] investigated the demographics of people with disabilities who knew about or had experience in crowd work. The research showed that some people with ASD use crowdsourcing already. However, the study did not reveal the challenges that this population face in working on micro-tasks.

In this paper, we investigate the opportunities and challenges in introducing people with ASD to work in a crowdsourcing environment. Some aspects of crowd work, such as bypassing the social norms of a traditional workplace, may be beneficial for people with ASD. Our hope is to improve the independence of those who face vocational challenges. This has become more possible as some states in the U.S. (*e.g.,* Pennsylvania) have changed their laws so that people receiving Social Security Income can earn more supplemental income without losing their eligibility for benefits [1]. Crowd work, however, could also introduce new challenges, such as requiring persons with ASD to navigate an overwhelming variety of job opportunities to successfully secure work. Our research questions include: Does crowd work present a feasible job opportunity for individuals with severe ASD? If so, what types of tasks would be best suited? And, could technology be designed to facilitate this kind of work?

To evaluate the potential of crowd work for providing individuals with ASD with new job opportunities, we conducted an iterative user-centered design study with three individuals with ASD. We assessed how well they were able to negotiate the existing crowd work environment (we focused on Amazon Mechanical Turk or AMT) and whether we could design technologies to facilitate their completion of micro-tasks. The study involved six weekly sessions with three participants with ASD. We dedicated the first four sessions to developing a better understanding of the participants' capacities in using the existing AMT interface to search and perform micro-tasks. We designed the sessions with an expert consultant with ASD. Based on findings from this stage, we designed, developed, and evaluated a prototype tool—Assistive Task Queue (ATQ)—to support our participants while working on image transcription tasks. Although, it is impossible to generalize from three participants, even with only three participants we observed substantial diversity of abilities. Over six weeks, we were nevertheless able to iterate toward approaches that seemed to work well for each participant, suggesting broader applicability of our general approach.

Our findings from the in-lab observations and in-lab studies suggest that our participants' abilities to work on micro-tasks vary significantly. But all could perform some tasks. Our participants performed best on tasks that required a low degree of executive function. For example, they could complete image description and image transcription that required limited working memory and relied on visual perception. On the other hand, they struggled with

[1] PA ABLE. http://www.paable.gov

a writing task that taxed working memory. We also found different levels of ability to search and explore the micro-tasks available on AMT. One of our participants could independently search and find tasks to work on, while the other two needed guidance from a researcher. Our findings are complicated by comorbid cognitive impairments in all participants, although such a combination of disabilities is common in this population [2]. In evaluating ATQ, we showed that the tool could facilitate our participants with ASD to find and work on image transcription.

This work qualitatively and quantitatively showed that our participants with ASD can perform crowd work. And, they benefited from simplification and guidance of micro-tasks and help with automated task search. More broadly, we contribute: (i) the first formative evaluation of challenges in introducing people with ASD in a crowd work environment, (ii) the assessment of how different types of micro-tasks are suited for providing individuals with ASD with new job opportunities, and (iii) the design and evaluation of a ATQ to facilitate work on image transcription tasks for our participants with ASD.

2. Background and Related Work

We describe the challenges that traditional workplace settings pose to people with ASD. We also introduce the HCI research that focuses on supporting this population.

2.1 Employment for Adults with ASD

It is estimated that there are approximately 3.5 million people with ASD in the U.S. [1], of which 50-75% of adults are unemployed [15]. Unemployment not only affects independence and self-esteem, but also significantly impacts the economy. For instance, it is estimated that $1.4-2.4m is needed to support an individual with ASD over their lifespan [8, 23]. Even those who are considered higher functioning are less likely to be employed and make less money compared to people without ASD.

While the symptoms manifested by adults with ASD vary across individuals [36], interaction and behavioral difficulties negatively impact employment [15, 19]. Prior work reports that interaction disorders pose the largest difficulties for individual with ASD in the workplace [15]. The challenges include difficulties understanding directions, in face-to-face communications (*e.g.*, reading facial expressions), and communicating in an inappropriate manner. Furthermore, behavioral difficulties such as throwing tantrums, physical aggression, self-injury, and ritualistic behaviors (*e.g.*, stimming) can create employment barriers [10].

Cognitive dysfunction and comorbidity (*i.e.*, the simultaneous presence of two or more chronic conditions) with other psychiatric symptoms could also negatively affect job opportunities. Difficulties in task execution due to problems with attention, motor planning, and working memory likely negatively impact work performance [32]. Acclimating to new job routines and managing changes in the work setting often pose additional challenges (though it can be overcome [20]). The individuals with ASD are sometimes also diagnosed with symptoms like depression and anxiety [22], which requires additional care.

Vocational programs and societal changes have been improving the situation [11, 14–16]. The programs attempt to help individuals prepare for and gain employment through various interventions, including job placement support (*e.g.*, job matching), on-the-job support (*e.g.*, increasing the awareness of employer toward ASD), and workplace modifications (*e.g.*, reducing over stimulating distractions). More recently, major technology companies—namely Microsoft and SAP—has been hiring people with ASD because they may be well-suited for the types of jobs offered in the technology-industry. Survey and interview studies conducted by Morris *et al.* revealed that software engineering tasks that require attention to details (like software testing) seems to be well suited for adults with ASD. While valuable, skill-matching alone does not address challenges with securing/keeping jobs and the volume of jobs offered by the pioneering companies are limited. Clearly, more work in addition to these efforts is needed to improve the situation.

	Age	Spectrum	Comorbidity	Occupation	Living
P1	28	Autism disorder	Intellectual disability	Janitor	with parents
P2	23	Asperger's syndrome	ADHD; Intellectual disability; OCD; Schizophrenia	Janitor	with housemates
P3	22	Autism disorder	Intellectual disability	Packaging	with parents

Table 1. Participant demographics from the survey, interview, and conversation with the participants' caregivers.

The crowdsourcing environment could offer new work opportunities for people with ASD. For example, a work environment where communication is nonverbal and asynchronous in nature may be favored by people with ASD [15, 20]. At the same time, however, aspects of crowdsourcing such as the lack of on-site support and overwhelming diversity of types of work offered could prevent people with ASD from working on crowdsourcing markets. Crowdsourcing has been studied for more than a decade, but limited research has assessed the potential benefits and drawbacks of crowdsourcing for people with ASD.

2.2 HCI Research on ASD

A range of HCI research has investigated the roles that technology could play in supporting people with ASD and their caregivers. Researchers have studied online communication by people with ASD [9, 34], designed methods/technologies for education and intervention [25, 33], and designed technologies to support children and their caregivers [13, 21]. For example, Hayes *et al.* developed CareLog, a tool to support school teachers to better conduct functional behavior assessment [13]. For more comprehensive survey of HCI research on ASD, see the recent survey by Kientz *et al.* [21].

The above research focused on supporting children and caregivers, and limited work looked into vocational support for adults with ASD. Morris *et al.* surveyed a neuro-diverse population (including people with ASD) to understand the challenges and opportunities they find while working at a large technology company [30]. While related, the focus was not understanding how people with ASD use crowdsourcing. Hong *et al.* studied the design of a social network to support the independent living of adults with ASD [17]. However, it was about designing systems to let people with ASD learn life skills through online collaborations. Finally, Swaminathan *et al.* investigated web accessibility of the crowd work environment [38] and Zyskowski *et al.* investigated who works on AMT [42], but their focus was not on people with ASD.

3. Study Method

Our study used the iterative user-centered design inspired by the participatory design method [37], a process that uses continual participation of the target population. The method was uniquely suited for this study; iteratively conducted study sessions not only let us understand how our participants use AMT, but also allowed us to simultaneously envision and shape new tools [37]. Furthermore, the method has been previously applied to accessibility research [29, 40]. We used AMT because it is widely used both in HCI research and in practice, and it offers a variety of micro-tasks. We asked our participants to perform the types of tasks that are common in the crowdsourcing market, such as image tagging, image transcription, and survey tasks [7, 12, 38].

Working with a local research partner, we recruited three participants with ASD through referral (Table 1). Our participants had other disabilities, too (Table 1). Participants had prior experience in using a computer. The participants used the computer in our lab to work on tasks administered in the study. The participants were compensated $20 per session in addition to what they earned by completing Human Intelligence Tasks (HITs) on AMT. All study sessions were audio and video recorded and interviews were transcribed by the researchers. The participants' interactions were recorded using a screen recorder. In each session, one participant and participant's caregiver attended the meeting. The study lasted for five to six sessions per participant.

4. Iterative User-Centered Design

The study lasted for six weeks from March to May in 2017. The first four weeks were spent to explore the participants' abilities to search and work on micro-tasks. In the following week, we designed, developed, and evaluated an assistive tool to support people to perform image transcription tasks—task that is prevalent on AMT and could be performed by all of our participants.

We describe the study and the summary of the findings from each week. We also discuss how the findings affected the subsequent weeks' research directions. In designing the sessions, we got feedback from an expert consultant who also has high-functioning ASD. The in-lab studies necessarily used within-subjects design given the expected variability between our participants. We relied on researcher observations more often than think-aloud because our participants' cognitive and communication disabilities made it hard for them to fully articulate their thoughts.

4.1 Week 1: Introductory Session

4.1.1 Method

We administered a background survey to collect demographic information including name, age, sex, comorbidity with other disabilities, employment status, and experience with using computing devices. When needed, the researcher and/or caregiver explained the intent of the questions. We then introduced the basic mechanism of crowd work, AMT specific concepts (*e.g.,* HITs), how to interact with the AMT interface, and types of tasks available. In the interview, we asked about their experiences on the current/most recent occupation, interpersonal challenges they face, their preferences on work location/time flexibility, their expectations toward crowd work, and experience in remote work.

To see how people with ASD use the crowdsourcing platform, the participants were asked to freely explore and work on AMT tasks. We were interested in their task search behaviors, task selection behaviors (*e.g.,* types of tasks they select), and how they work on the tasks. When the participants struggled to find tasks to work on, we offered the following keywords for searching: *data collection*, (image/audio/video) *transcription*, (image/video) *description*, *survey*, and *writing*. These suggestions were made based on the types of task that are common on AMT [7, 12]. If the participants further struggled, we ask them to work on the micro-tasks of the above categories that we made and posted on AMT.

4.1.2 Result

P1 and P2 participated from Week 1. P3 participated from Week 2 but went through the Week 1 procedure except for working on AMT. The sessions with P1 and P2 took 64 and 72 minutes. All participants were male. The age ranged from 22 to 28. P1 and P3 had autism disorders, and P2 had Asperger's syndrome. All participants had comorbidity (*e.g.,* intellectual disabilities—see Table 1). All participants found their current or most recent occupations through our research partner's referral. P2 earned $77-130 per week; P1 and P3 did not know their weekly earnings.

I would be interested in crowd work because I could	**P1**	**P2**	**P3***
earn money	4	4	5
do something interesting or fun	4	5	5
do something that's stimulating or challenging	4	4	5
learn new skills	4	5	5
feel like I'm contributing to the society	3	5	5
help other people on their projects	3	5	5
participate with family and friends	4	5	5
work without interpersonal communication	3	5	5
work anonymously	2	3	5
work remotely at anywhere I want	4	4	5
work at anytime.	2	3	5

Table 2. Participants' motivation for working on crowdsourcing. The scale ranged from 1 (strongly disagree) to 5 (strongly agree). *The response from P3 are from Week 2.

When asked about their work preferences, P1 and P2 said they like routine jobs with fixed schedules. None had prior experience in remote work. They had no knowledge about crowdsourcing until they heard about the study. They had no preference whether the work was on-site or remote. We asked to rate why they were interested in crowd work in 5-point scale—see Table 2. When asked if they want to treat crowd work as either main source of income or supplementary source of income, all responded that they want to work on crowd work in addition to other jobs.

P1 and P2 were asked to freely work on AMT tasks. We observed their varying abilities to search and work on micro tasks (*i.e.,* HITs). P1 who has an autism disorder required extensive help and guidance from the researcher. The researcher provided step-by-step instructions on searching for HITs and prompted him to work on them. Completing a HIT was also difficult. For example, when P1 worked on an image description task where he was asked to provide a title and an alt-text for a given image, he had trouble focusing on what to describe. Seeing the struggle, the researcher decided to ask P1 to work on an image classification task that we prepared, where the participant was asked to confirm the presence of a target object in a series of pictures, which he could complete. P2, on the other hand, could independently use the AMT interface and find HITs that he could work with minimal guidance.

4.1.3 Week 1 Summary

In general, our participants were motivated to work on AMT tasks. The factors like *earning money, do something interesting or fun* and *stimulating, learning new skills,* and being able to work *with caregiver* seemed to be important for all participants. *Anonymity* and *work time flexibility* seemed to matter less for P1 and P2. Although they had no preferences on on-site or remote work, they were still interested in the remoteness of crowd work.

The abilities of P1 and P2 to work on micro-tasks varied. P1 struggled on searching and completing HITs, whereas P2 had no problem navigating the AMT interface. It is important to note that both participants could work on some HITs, suggesting the feasibility of introducing adults with ASD to work on micro-tasks. We further investigated what kinds of micro tasks they could work on in Week 2 and 3. We also studied how we could support them to find HITs in Week 4.

4.2 Week 2: Task Completion Efficacy 1

To better understand our participants' abilities to work on micro-tasks, we asked them to work on a variety of micro-tasks. P3

Figure 1. The examples of images used in the Week 2 tasks. (a) *Image classification*: prompted yes/no question like "is this this woman smiling?" (b) *Image description*: prompted to describe what is in the image. (c) *Image transcription*: prompted to transcribe parts of the image. like the first name in the business card.

joined and he went through the Week 1 procedure, except for the part where we asked to freely explore AMT. Overall, the sessions with P1, P2, and P3 took 68, 67, and 108 minutes.

4.2.1 Method

We prepared the following five categories of tasks: (i) *image classification*, (ii) *image description*, (iii) *image transcription*, (iv) *survey*, and (v) *writing*. The goal was to prepare tasks of varying types and difficulty levels that exist on AMT. For image classification and image description tasks, we downloaded 5 images for each from ImageNet that were previously tagged by AMT workers (Figure 1a). The image classification task prompted the participant to answer five yes/no questions like "is this woman smiling?" The image description task asked the participants to describe and enter what is in each image (Figure 1b).

For the image transcription, survey, and writing tasks, we downloaded the HITs' HTML files and media contents that were available on AMT at the time, and re-posted them as our HITs. We had two images for the image transcription task: a shopping receipt and a business card (Figure 1c). The participants were asked to find and transcribe 12 pieces of information like "total amount of transaction in the receipt" and "email address in the business card." The survey task asked to score the relevancy of 9 search results for a given term "Frozen Yogurt". The query results included items like "Dawns Deli" and "YoGo Factory Frozen Yogurt" and the participant were asked to provide 5-point Likert response that ranged from highly relevant to highly irrelevant.

In the writing task, participants received two lists of information about two imaginary restaurants. Each list included restaurant's name, location and cuisine, as well as the qualities of decor, food, and service (*e.g.,* "Neighborhood is Palo Alto", "Service is mediocre."). The participants were asked to generate a fluent description that compares the two restaurants. All the tasks were posted on AMT. We compensated the participants with $0.05-0.10 per task (in addition to the $20 session pay). In addition to *in-situ* observation, we report the time they took to complete each task as a proxy of how hard the tasks were for the participants.

4.2.2 Result

The participants spent 28 seconds to 21.6 minutes to complete the types of the tasks that we prepared (Table 3).

	Image Classification	Image Description	Image Transcription	Survey	Writing
P1	1.72	3.68	10.75	7.40	21.62
P2	0.47	0.73	2.63	1.08	7.95
P3	6.17	3.40	19.87	3.72	10.88

Table 3. Task completion time (min) for each task category. The participants finished classification and description tasks relatively quickly. P1 and P3 took long time to complete transcription and writing tasks.

Image Classification. Image classification tasks took the least amount of time for P1 and P2, but P3 took longer than the others. While P2 had no problem, P1 and P3 required guidance in answering the questions; the researcher needed to prompt for an answer to each question (*e.g.*, asking "is this a picture of soccer game?"). P3 also had challenge in providing accurate classifications. He had tendency to answer "yes" to all the classification questions. Note that he knew the correct answer, but the answer was not correctly registered. For example, for one question with a picture of people playing soccer that asked "is this a picture of people playing baseball?" P2 correctly said it is soccer, but he still selected "yes" for the answer.

Image Description. The image description task took a moderate amount of time for the participants. All participants provided accurate descriptions for the pictures. Like in the image classification task, the researcher needed to initiate each question for P1 and P3 by asking "what is in the next image?"

Image Transcription. P2 completed the image transcription without guidance, but P1 and P3 struggled to initiate transcription, like in previous tasks. P1 and P3 took a long time to complete the task because searching for the right information in the images was hard. In the name card transcription, for instance, the researcher needed to provide step-by-step guidance like "is the company name on the image?", "where do you see it?", "what is the company name?", and "can you transcribe that?"

Survey. P1 and P3 struggled to complete the task. It seemed that the task was more taxing for P1; he took two times longer than P3 to complete it and four times more than the time he spent on image classification. P2 had no problem completing the task.

Writing. The writing task took the most time for P1 and P2. It took a significant amount of time for P3, too. Constructing natural sentences while using the given lists of information seemed taxing for the participants. The generated text were also not fluent. For example, P3 wrote *"restaurant are lemongrass and benjarong in palo alto. cuisine is thai. decor is excellent. food quality is excellent. service is mediocre and good,"* which is almost an exact copy of the information provided on the lists.

4.2.3 Week 2 Summary

The abilities to complete the tasks varied across the participants; P2 could perform most of the tasks that we prepared, but P1 and P3 struggled. The participants could perform image description and survey tasks although P1 and P3 needed help starting to answer. Image classification could be performed relatively easily by all, but P3 had a unique challenge in registering what he observed in the given picture to the yes/no answer. This could mean that he was fixated on answering "yes" to all the questions, he did not understand the question, or both.

The tasks that required too much information processing seemed to be hard. All participants struggled to perform the writing task. It seemed challenging to keep the given list of information in short-term memory and synthesize it to generate text. This observation could be explained by the current understanding of executive dysfunction in people ASD. That is, writing tasks such as what we administered tax working memory [28].

In the image transcription task, reading and transcribing seemed doable for our participants, but finding the right information on the image seemed hard for P1 and P3. It is not clear, however, whether this is the side-effect of executive dysfunction (*i.e.*, challenge to plan and systematically search for the information to transcribe) or visual attention bias. Since P2 seems to be unaffected, we think it is not common across all people with ASD.

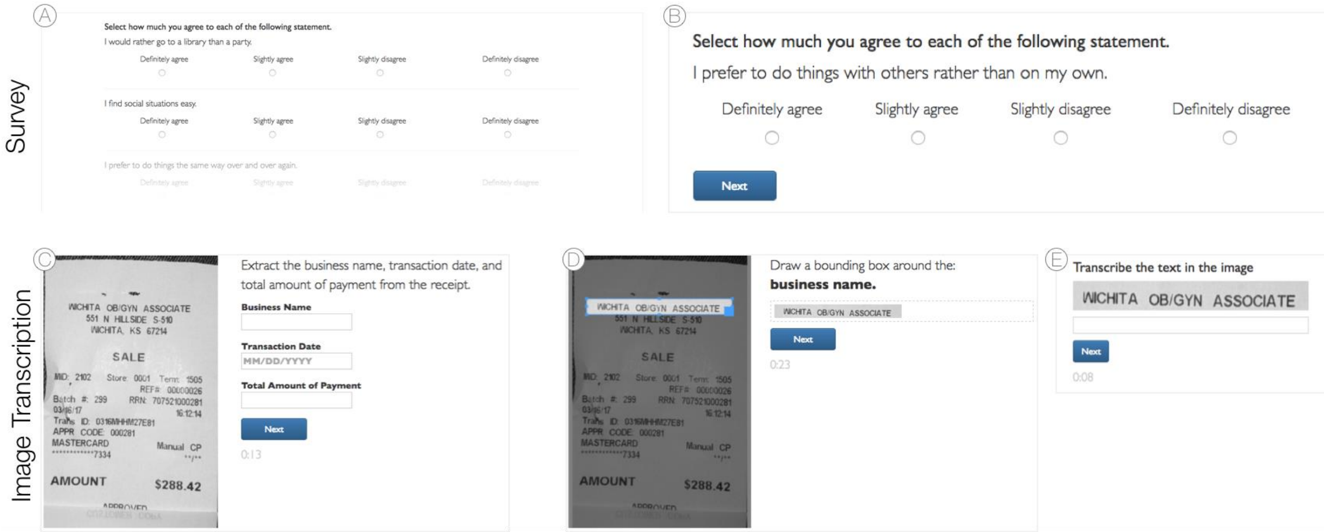

Figure 2. The task interfaces for Week 3. (a) *Survey control*: 10 questions from AQ were listed on a single page. (b) *Survey mini*: 10 questions from AQ were administered one-by-one to the participant. (c) *Transcription control*: we asked the participant to provide business name, transaction date, and total amount of payment for each receipt in a single page. (d) *Text detection*: the participant used the interface to draw a bounding box around the target information like business name. (e) *Transcription mini*: The participants were asked to transcribe what was on the cropped image of the receipts.

4.3 Week 3: Task Completion Efficacy 2

Given the results from Week 2, we decided to further investigate whether our participants could perform image transcription tasks. We also investigated the performance on the survey task as it was not clear how well our participants performed on these tasks. We also decided to administer these tasks due to their increasing prevalence on AMT [12, 38]. Overall, the sessions with P1, P2, and P3 took 62, 37, and 59 minutes respectively.

4.3.1 Method

In addition to assessing our participants' ability to work on image transcription and survey tasks, we also investigated the potential effect of task simplification. Both survey and transcription tasks had control and simplified interface conditions. We measured accuracy and task completion time.

Survey. We used questions from the Autism-Spectrum Quotient (AQ) [3]. AQ was used primarily because of its tested internal consistency of the survey items and secondarily because we wanted to see how our participants respond to the questions. We had two conditions: *control* interface that has the questions in a single page (Figure 2a) and a simplified *mini* interface where the questions were served one-by-one. Each condition had a set of 10 questions. Each set took a pair of two questions from the five different areas of constructs in the AQ—social skill, attention switching, attention detail, communication, and imagination—and each pair consisted one "agree" response and "disagree" response [3]. This was done to test the repeatability.

Image Transcription. In preparing the image transcription tasks, we downloaded images from receipt transcription HITs that were available on AMT. The task involved transcribing business names, transaction dates, and total transaction amounts from 10 receipts. We tested two methods: (i) a *control transcription* condition in which we showed the image transcription interface that replicated the original HIT (Figure 2c), and (ii) a simplified *mini transcription* condition where the participants were asked to transcribe string in each image where the target area was cropped (Figure 2e). The mini transcription condition therefore reduced the burden of searching for the right information to transcribe. In addition, we asked them to work on *text detection* in which we asked them to draw bounding box around the target information (Figure 2d). Our goal was to see the time cost and the accuracy of visual search. For each condition, there were two practice tasks.

	Survey Control	Survey Mini	Transcription Control	Transcription Mini	Text Detection
P1	0.6 (3/5)	0.0 (0/5)	1.0 (30/30)	1.0 (30/30)	1.0 (30/30)
P2	0.2 (1/5)	0.2 (1/5)	0.9 (27/30)	0.97 (29/30)	0.77 (23/30)
P3	0.2 (1/5)	0.0 (0/5)	0.8 (24/30)	1.0 (30/30)	0.8 (24/30)

Table 4. Task accuracy for the Week 3 survey tasks and transcription tasks. For survey control and mini conditions, we used repeatability as a proxy for the accuracy.

Measures. For image transcription tasks, we measured task completion time and accuracy. Transcribed strings were compared to researcher-prepared ground truth transcription to measure the accuracy. Special characters (*e.g.*, '$') were omitted from the analysis and the date format was disregarded (*e.g.*, "Mar 29 '17" and "29 03 2017" were treated as same). Accuracy for the bounding boxes in text detection tasks were visually assessed by a researcher. Because transcribing three items (business name, transaction date, and transaction amount) were split in the *mini transcription* and *text detection*, we added the time in each triplet to compare with the *control* condition. Similarly, we measured task completion time and accuracy for the survey tasks. Because there is no "correct" answer for surveys in general, we use the repeatability as a proxy for accuracy. Repeatability was measured by counting agreements of responses to the items in the same construct (*e.g.*, items "I find social situations easy" and "I would rather go to a library than a party" are in the same construct *social skill* that have opposite polarity). For example, if responses in 3 out of 5 constructs agree, then the overall agreement is 60%.

4.3.2 Accuracy

Survey. The repeatability measures were 0.6, 0.2, and 0.2 for the *control* tasks and 0, 0.2, 0 for *mini* tasks for P1, P2, and P3 respectively (Table 4). The responses from P2 and P3 had low repeatability in both conditions. Interestingly, the response from P1 had lower repeatability in the *mini* condition. The researcher who observed the session noted that P1 seemed to be fixated on choosing the same answer on the *mini* condition rather than reading and answering the questions.

***Transcription*.** P1, P2, and P3 transcribed the images with 100%, 90%, and 80% accuracy in the *control* condition and 100%, 97%, and 100% in the *mini* condition (Table 4). P1 transcribed text perfectly in both conditions. P2 made typos when entering total transaction amount, which seemed more prevalent in *control* interface. P3 made 6 errors, of which 5 were mistakes where he entered addresses of the businesses instead of the business names. To compare the transcription accuracies between *control* and *mini*, we used generalized linear mixed model (GLMM) [4]. We used GLMM instead of the oft-used repeated-measures ANOVA to incorporate the random effect (*i.e.*, subject variability). As the responses were binary (correct/incorrect), we used a logit link function. We had interface (*i.e.*, *control vs. mini*) as the main effect, and we had the intercepts for the participants as the random effect. There was a significant main effect of the interface ($\chi^2(1, N{=}180){=}7.9$; $p{<}0.01$), showing that simplifying tasks had a significant positive effect on the accuracy of the task.

Text detection accuracies varied across participants. P1 performed the task perfectly, but P2 and P3 made some mistakes (Table 4). The source of P2's errors were bounding boxes that contained not only the target information like business name, but also other things like business addresses.

4.3.3 Task Completion Time

***Survey*.** On average, P1, P2, and P3 took 21.6, 4.5, and 23.1 seconds to answer items in the survey *control* condition, and 11.7, 5.3, and 15.3 seconds to complete in the survey *mini* condition. Figure 3 shows that P1 and P3 took less time to answer questions on the *mini* condition, but it seems that there was a limited effect for P2. To test the significance, we used a linear mixed model (LMM). Like the accuracy analysis, we had *interface* as the main effect and *participants* as the random effect. We observed the significant main effect of the interface type ($\chi^2(1, N{=}60){=}9.65$; $p{<}0.01$). This shows that simplifying the survey task made task completion significantly faster. Note, however, the speed up may be due to P1 getting fixated on repeating (*i.e.*, not reading and comprehending the questions) rather than task simplification.

***Transcription*.** On average, P1, P2, and P3 took 64.9, 27.4, and 57.6 seconds to transcribe a receipt in the *control* condition, 45.4, 24.6, and 37.0 seconds in the *mini* condition. Text detection took 90.2, 41.4, and 74.6 seconds on average. Looking at Figure 3, participants, especially P1 and P3, seem to have transcribed images faster in the *mini* condition compared to the *control* condition. We used LMM to test the significance of the effect of interfaces. We observed a significant main effect of interface ($\chi^2(1, N{=}60){=}21.2$; $p{<}0.001$). This shows that removing the visual search component of the transcription task significantly reduces the task completion time. Note, however, making mini interface requires text to be detected and cropped (Figure 3). When our participants worked on text detection, it took significantly longer time than the *control* transcription ($\chi^2(1, N{=}60){=}17.6$; $p{<}0.001$) due to the added cost of the image cropping interaction.

4.3.4 Week 3 Summary

The participants finished survey tasks faster due to task simplification. However, breaking the survey into smaller pieces had an unexpected negative impact on the participants' response quality. For instance, P1 selected "slightly agree" repeatedly to all the questions in the *mini* condition. We asked why he chose the same answer for every question, but he could not articulate the reason. The speed-quality tradeoff makes it unclear whether it is beneficial to break down the survey only to make it easier for the participants with ASD. While we believe the low repeatability was because of the participants' limited abilities to comprehend the survey questions, we cannot strongly argue this was not dues to the insufficient internal consistency of the survey questions as we used a subset of the items in AQ (instead of the original AQ).

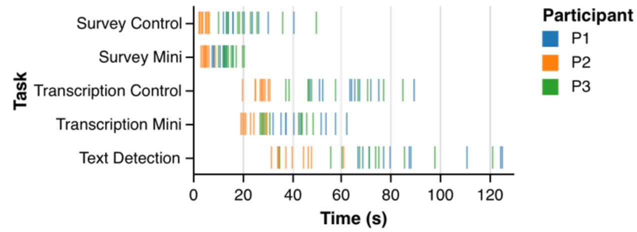

Figure 3. The strip plot of task completion time by each participant. The x-axis is completion time in seconds and y-axis is the category of the task.

We observed a significant positive effect of task simplification on both quality and speed for image transcription. The improvement in the task speed is likely due to the reduced burden on searching for relevant information to transcribe. We were somewhat surprised to see the positive effect on the transcription accuracy for P2 and P3. Reduced task complexity that allowed people to focus on what to transcribe may have influenced the accuracy. However, this could also be a learning effect as the *mini* condition was administered after the *control* condition.

Reducing task complexity itself takes time. For example, *text detection* alone took significantly longer than *transcription control*. Note that breaking tasks down into smaller pieces has the added benefit of giving more flexibility on how to split the task between workers. Thus, for example, those who can detect text could work on text detection, and those who are good at transcribing could focus on text transcription. *Subcontracting microwork* [31] could facilitate the process. Future work should further investigate the cost-benefit tradeoff of task simplification.

4.4 Week 4: Task Search Behavior

We observed that searching for HITs introduced added interaction complexity and reduced the productivity of P1. To further investigate how our participants search and explore HITs, we asked them to search for image transcription tasks using the AMT interface and observed their behaviors.

4.4.1 Method

We asked our participants to find image transcription tasks using the AMT's default search interface. We presented the top page of the AMT's worker page, then asked them to use the keyword "transcribe" to find image transcription tasks using the AMT interface. We asked them to complete five transcription tasks.

***Measures*.** In addition to the *in-situ* observation, we counted the number of the tasks they completed during the approximately hour-long session, the number of times that they opened the non-image transcription tasks (*e.g.*, audio/video transcription tasks), how long it took to find the task to work on, and how long it took to complete each task. We did not measure transcription accuracy because the tasks differed across the participants.

4.4.2 Result

Our participants spent, on average, 9.9% of their time on the task search. P1, P2, and P3 completed 2, 5, and 4 tasks respectively.

P1: During the 68.2 minutes session, P1 found and completed two image transcription tasks. He spent 5.9 minutes for task search—8.6% of the session. Finding HITs to work on required support from the researchers as the participant could not figure out how to navigate AMT's search interface. Completing two tasks took more than an hour. The tasks that he worked on involved transcribing scanned images of hand written text. Decoding

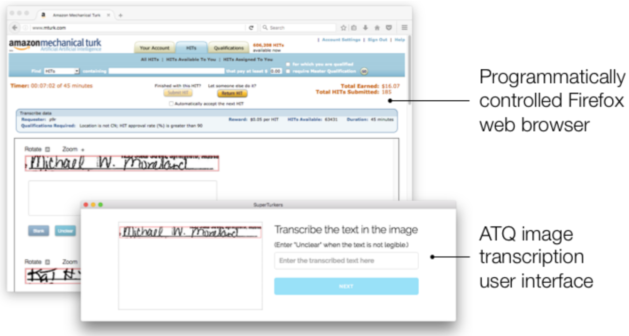

Figure 4. The ATQ prototype. The system has a custom desktop UI for image transcription, a programmatically controlled Firefox web browser, and a Selenium browser automation tool that runs on the background.

unclear handwriting seemed more challenging compared to transcribing printed texts in receipts or business cards (which he did in Week 2 and 3). P1 often got stuck when he could not read unclear text and the researcher intervened. For example, the researcher instructed to enter "unclear" if the text is not legible.

P2: P2 found and completed 5 transcription tasks in 20.1 minutes. He spent 1.3 minutes for searching the tasks (6.4% of the time to complete 5 transcription tasks). For both task search and working on image transcription, he did not require help from the researcher.

P3: P3 completed 4 transcription tasks in 59 minutes. He spent 8.6 minutes searching for the tasks and 50.2 minutes for completing the HITs. Task search took 14.6% of the session. While searching for the image transcription tasks, he opened two audio transcription tasks. Because he did not know how to go back and search for image transcription tasks, the researcher intervened and instructed how to resume the search. While working on HITs, P3 did not get stuck but often made errors.

4.4.3 Week 4 Summary

The observations from Week 4 revealed that our participants spent a non-negligible amount of time looking for HITs. For example, while searching for a task to work on, P3 clicked and opened two audio transcription tasks. Because the goal was to find and work on image transcription tasks, we asked him to search for another task in those cases. Note, while this restriction was artificial (*i.e.*, he may be able to work on audio transcription), it is not unrealistic that he will encounter many tasks that he cannot perform if he were to freely use AMT. It appears crucial to reduce worker task search time, which is unpaid work.

Our findings from Week 1 to 4 pointed to two potential research directions. First, because abilities of people with ASD and skills required to perform certain HITs both vary widely, it seems that technologies that match people and tasks based on both of these features seem helpful. Second, we believe technologies that semi-automatically simplify searching and working on micro-tasks will improve the productivity of people with ASD. In the rest of the paper, we focus on the latter and leave the former as future work.

4.5 Assistive Task Queue

We designed and developed a prototype Assistive Task Queue (ATQ) that simplifies working on image transcription micro-tasks for our participants with ASD. ATQ was developed with three design principles: (i) provide tasks that can be performed by the participants with ASD, (ii) remove the complexity of searching for HITs, and (iii) breaking a large task into smaller pieces that are more digestible for the participants with ASD.

First, given the limited abilities of our participants with ASD to perform AMT tasks, feeding them the tasks that they can perform is likely important for them to successfully complete the HITs. Based on our findings from Week 1 to 4, we believe that tasks that do not tax working memory and executive function in general would be good candidates for our participants. For example, image description, image classification, and image transcription that relies more on perception and less on comprehension and synthesis of ideas seem more feasible. This observation and the popularity of image transcription tasks on AMT [12] made us to focus on designing tools to facilitate image transcription tasks.

Second, we design ATQ to reduce the time for task search. ATQ automatically crawls, finds, and accepts image transcription tasks on behalf of its user. Third, ATQ breaks down the web interface of an image transcription HIT into smaller pieces and feeds them one-by-one to its user.

We use the following components to implement the prototype ATQ: the custom desktop transcription UI developed with HTML, CSS, and JavaScript, Firefox web browser, and Selenium browser automation tool (Figure 4; Video Figure). When the user opens the ATQ, Selenium requests the browser to search for a list of available image transcription HITs on AMT using predefined search queries. Once HITs appear on the browser window, Selenium parses the HIT titles and match them with predefined titles of image transcription tasks. The first task that matches gets programmatically accepted. Once the HTML content of the accepted HIT appears on the browser window, the content gets sent to the transcription interface by Selenium and it gets parsed. We use a predefined pattern for each HIT to parse its HTML.

Based on the parsed information, the image transcription interface shows an image to transcribe as well as the input field. One image is presented at a time. The transcription interface then prompts the user to transcribe the shown image like what the original interface asks the user to do. Once the user types in the information to the input field and click "Next" or hit enter, the transcribed information is sent to the Firefox browser and fills the corresponding input field on the original HIT. Once all the input fields were entered, the HIT was automatically submitted.

4.6 Week 5: ATQ Evaluation

We evaluated the usability with the three participants to investigate the utility of ATQ. Overall, the sessions with P1, P2, and P3 took 48, 31, and 44 minutes.

4.6.1 Method

Procedure. The participants were asked to work on transcribing text from a series of scanned images, each containing a line of hand written text (Figure 4). There were two conditions: the *control* condition that used the original interface that was posted on AMT and the *ATQ* condition. We downloaded the image transcription tasks posted by a requester *p9r* and re-posted them as our tasks. This was necessary to administer the same images across the participants. In both conditions, the participants were asked to transcribe twelve practice images and twenty images. The images included strings like names, email addresses, and medical conditions. The tasks were served in the following order: practice control task, control task, practice ATQ task, and ATQ task. At the beginning of each task, we instructed the participant to start transcribing the image. We measured the time for transcribing each image. Once a participant started a task, we intervened as little as possible, but if he was stuck for more than 10 seconds, we gave instructions (*e.g.,* we told to type "unclear" and move on when the text on an image was not clear).

Transcription Accuracy						
	Control (N=20)			ATQ (N=20)		
	Correct	Incorrect	Unclear	Correct	Incorrect	Unclear
P1	9	3	8	17	3	0
P2	17	3	0	16	2	2
P3	12	8	0	14	6	0

Table 7. The numbers of correct/incorrect transcription responses. Unclear indicates how many times the participant noted the text is not legible.

Measures. In addition to in-situ observation, we looked into the transcription accuracies and interaction intervals. A member of the research team transcribed all the images and matched them with the participants' transcripts to assess the accuracy. Up to two character typos were forgiven. Interaction intervals (*e.g.*, how long each participant spent typing) and total task completion time were measured from the screen recording.

4.6.2 Accuracy

P1, P2, and P3 transcribed 9, 17, and 12 out of 20 images correctly in the *control* condition, and 17, 16, and 14 out of 20 correctly in the *ATQ* condition. P1 and P3 transcribed text more accurately using ATQ, while P2 performed equally correctly in the both conditions (Table 7). We used GLMM with a logit link function to assess the significance. We had *interface* as the main effect. We had a random intercept for participants. We did not observe a significant main effect of interface, but there was a possible trend ($\chi^2(1, N=120)=3.74$; $p=0.053<0.1$). This suggests that more work is needed to investigate whether the interface condition affected the transcription accuracy.

4.6.3 Task Completion Time.

P1, P2, and P3 took 13.2, 4.9, and 9.5 minutes respectively to complete the task on the *control* condition. All participants finished the task faster in the *ATQ* condition; P1, P2, and P3 took 10.0, 3.4, and 8.8 minutes respectively (Figure 5; Table 5). We used LMM to evaluate the significance of the difference between the conditions. We had *interface* as the main effect and *participant* as the random effect. A significant main effect of *interface* was observed ($\chi^2(1, N=120)=4.02$; $p<0.05$). Faster task completion time in the *ATQ* condition was not due to the faster typing speed, but because of the reduced interaction complexity. For example, ATQ condition did not require the participant to scroll the browser window to move to the next transcription image.

P1: P1 required more interventions than the others during the control task; the researcher suggested to enter "unclear" when he was stuck and paused for more than 10 seconds. We intervened 8 times in the control condition. He also made consistent mistakes on transcribing email addresses, where he usually skipped what was after "@". He consistently substituted special characters like "-" with spaces.

P2: P2 completed the tasks in the both conditions without problems (except for some typos), thus we did not intervene. He told us that ATQ condition "made the task more competitive," probably because it gave more game-like feeling (*e.g.*, task images showed up one after the other quickly).

P3: P3 completed all the task both in AMT interface and ATQ without any interventions. He made some errors in transcription.

4.6.4 Week 5 Summary

We evaluated the utility of ATQ to support our participants with ASD on image transcription. We observed a significant improvement in task completion time due to reduced interaction complexity. We also note that this speed up is in addition to the benefit of reducing the task search time that ATQ automatically takes care of. There was no significant main effect of interface type to transcription accuracy.

Transcription Task Completion Time (minutes)						
	Control			ATQ		
	Overall	Typing	Other	Overall	Typing	Other
P1	13.2	10.7	2.5	10	8.2	1.8
P2	4.9	3.3	1.6	3.4	2.9	0.5
P3	9.5	8.1	1.4	8.8	7.6	1.2

Table 5. Task completion time. Both conditions (*control* and *ATQ*) involved transcribing 20 images of text.

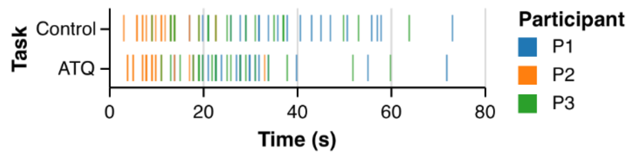

Figure 5. A strip plot of time spent on transcribing each image. P1 and P3 took longer to complete transcription in the *control* condition.

I would be interested in crowd work because I could	P1	P2	P3
earn money	5	5	4
do something interesting or fun	5	4	4
do something that's stimulating or challenging	5	4	4
learn new skills	4	5	4
feel like I'm contributing to the society	4	4	5
help other people on their projects	4	5	4
participate with family and friends	4	4	5
work without interpersonal communication	4	4	4
work anonymously	3	4	5
work remotely at anywhere I want	4	4	4
work at anytime.	4	4	5

Table 6. Participants' motivation for working on crowdsourcing in the future. The responses were measured during the exit interview. The scale ranged from 1 (strongly disagree) to 5 (strongly agree).

We also identified potential improvements that we could make to ATQ. For example, we could build task dependent instruction to reduce some systematic errors. For example, we may be able to prevent a user from forgetting to transcribe domain names in email addresses by prompting them to transcribe the entire text in the image. Limitations of the prototype are: it focused on searching and parsing predefined tasks on the market. The researcher programmed what tasks to search and how to parse the task, which will not scale in the real-life setting. Future work should investigate potential methods to populate variety of tasks that can be fed into ATQ.

4.7 Week 6: Exit Interview

We conducted exit interviews with our participants. We investigated if their motivation to work on crowdsourcing changed over the period of this study. The sessions took 32, 18, and 27 minutes for P1, P2, and P3 respectively.

Method. We conducted a structured interview. The same set of questions from the interview in Week 1 were used, but were phrased to ask if participants are interested in working on AMT after working on the crowd work during the study. We also asked questions like "are you satisfied with what you made by working on AMT tasks," "how much time do you want to spend on crowd work per week," and "do you use AMT outside of the study?"

Result. Overall, our participants were motivated to work on AMT in the future (Table 6). This did not seem to change significantly from Week 1 (Table 1). P1, P2, and P3 made \$1.17, \$3.00, and \$1.03 by completing HITs over the study sessions. P1 and P2 told us they would like to make more than what they earned. For example, P1 noted that he would like to make minimum wage. None of our participants used AMT outside of the study. P1 and P2 were still going through the process of making their own accounts, which took a long time because Amazon requested personal information to set up the payment accounts, which meant our research partner had to help them make their AMT accounts. P3 created his own AMT account, but he did not use it.

5. Discussion

All of our participants with ASD were able to successfully work on Amazon Mechanical Turk. The tasks that each individual could best perform varied between people, depending in part on the how their ASD and other disabilities manifest. This research was motivated by the premise that the crowdsourcing could provide people with ASD an alternative option for work environment. Crowd work seems to have traits that may benefit this population. For example, interaction difficulties associated with ASD that account for the biggest vocational impact [15] could be mitigated by the non-verbal nature of online crowd work. This aligns with the findings from prior work [5, 27] that reported people with ASD felt more control in social communication over textual media (*e.g.,* chat and forum). Although the types of tasks that can be performed by our participants varied, this study suggests that there are some micro-tasks that people with ASD can perform. Assistive tools like ATQ could improve their productivity, too.

So, *is crowd work good for people with ASD?* We are confident that it is feasible to introduce many people with ASD to crowd work given our experiences with our participants, but a few challenges need to be overcome so that we can confidently suggest them to work on this environment. The challenges include *effective job matching, scalable work guidance,* and *hourly wage*.

Worker-Job Matching. Many job options were overwhelming to our participants. As mentioned in the Week 4 summary, we believe that technologies that match people and tasks based on users' abilities and skills required to perform tasks would be helpful and necessary. Such technologies would need to: efficiently assess the abilities that the user possess; automatically tease out what skills are needed to perform certain HITs; and optimally match the user with the HITs based on the extracted information. This also aligns with prior research in ASD and the ethos of ability-based design [25, 41]. Future work includes the design, implementation, and evaluation of such mechanisms.

Scalable Work Guidance. We showed that ATQ could improve the productivity of image transcription tasks. But the current prototype relied on a predefined search query and an HTML parsing format. Future work should investigate how to scalably search and identify tasks that could be broken down. The types of work that ATQ supports should be expanded, too. This is important; although the volume of image transcription tasks that are available on AMT was high at the time of the study, the number of jobs available on the market could diminish.

Low Hourly Wage. The low hourly wage of crowd work has been discussed extensively in the prior work [18, 26]. It is difficult to earn the U.S. minimum wage currently on AMT. For example, Horton and Chilton reported that median wage is $1.38 per hour [18], only a fraction of the U.S. federal minimum wage ($7.25). The problem would be exacerbated for our participants, who performed tasks more slowly than workers without ASD. That said, experts on AMT claim hourly wages of over $10; tools that make workers for efficient may help workers in general.

6. Limitations

A limitation of our work is that only three people with ASD participated, with the tradeoff that participants were observed over multiple sessions. Opportunities for future work include a larger user study with people with ASD. We solely focused on the capacity of people with ASD to perform crowd work and did not explore the roles that caregivers could play. We plan to explore how they could support people with ASD to perform micro-tasks. Our participants had comorbidity (*e.g.,* intellectual disability, ADHD). This may have affected our observations and make it hard for us to argue all the findings are due to ASD. Nevertheless, studying this population is important as comorbidity is common in people with ASD (*e.g.,* the CDC estimates that 56% of children with ASD have below average intellectual abilities [2]). We did not exhaustively test the types of the micro-tasks available on AMT. The tasks were often selected by the researchers and not by the participants. This was necessary for the in-lab studies. More work is needed to explore what types of tasks they can best perform.

7. Conclusion

We conducted an iterative user-centered design study to investigate the challenges and opportunities in introducing people with ASD to crowdsourcing work environment. The process was central to understanding the abilities of people with ASD to perform crowd work and simultaneously design technology to support the target population. We found that our three participants had varying levels of abilities to work on micro-tasks. We believe that the tasks such as image description and image transcription that do not tax their executive function are easier for them to perform, while highly cognitive tasks like writing pose challenges. This work qualitatively and quantitatively showed that our participants with ASD could perform crowd work and they benefited from simplification and guidance of micro-tasks. Based on the findings, we designed and evaluated an assistive tool to improve the productivity of people with ASD to perform image transcription, a task type prevalent on crowdsourcing market.

8. ACKNOWLEDGMENTS

We thank Sarah Lehman for her feedback. We also thank Liz Carter and Jill Lehman for their advice in structuring our research. This work was developed under grants from the National Institute on Disability, Independent Living, and Rehabilitation Research (NIDILRR grant numbers 90DP00610100 and H133E140039).

9. REFERENCES

[1] Autism Costs More Than $2 Million Over Patient's Life: 2014. *https://www.bloomberg.com/news/articles/2014-06-09/autism-costs-more-than-2-million-over-patient-s-life*.

[2] Autism Spectrum Disorder - Data & Statistics: Prevalence: 2016. *https://www.cdc.gov/ncbddd/autism/data.html*.

[3] Baron-Cohen, S., Wheelwright, S., Skinner, R., Martin, J. and Clubley, E. 2001. The Autism-Spectrum Quotient (AQ): Evidence from Asperger Syndrome/High-Functioning Autism, Malesand Females, Scientists and Mathematicians. *Journal of Autism and Developmental Disorders*.

[4] Bates, D., Maechler, M., Bolker, B., Walker, S. and others 2014. *lme4: Linear mixed-effects models using Eigen and S4*.

[5] Benford, P. and Standen, P. 2009. The internet: a comfortable communication medium for people with Asperger syndrome (AS) and high functioning autism (HFA)? *Journal of Assistive Technologies*.

[6] Brault, M.W. 2012. *Americans With Disabilities: 2010, Household Economics Studies*.

[7] Brewer, R., Morris, M.R. and Piper, A.M. 2016. “Why Would Anybody Do This?”: Understanding Older Adults’ Motivations and Challenges in Crowd Work. *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*.
[8] Buescher, A.V.S., Cidav, Z., Knapp, M. and Mandell, D.S. 2014. Costs of autism spectrum disorders in the united kingdom and the united states. *JAMA Pediatrics*.
[9] Burke, M., Kraut, R. and Williams, D. 2010. Social Use of Computer-mediated Communication by Adults on the Autism Spectrum. *Proceedings of the 2010 ACM Conference on Computer Supported Cooperative Work*.
[10] Burt, D.B., Fuller, S.P. and Lewis, K.R. 1991. Brief report: Competitive employment of adults with autism. *Journal of Autism and Developmental Disorders*.
[11] Cullum, P.M. and Ennis-Cole, D. 2014. Preparing adults with autism spectrum disorders for employment. *Online Journal for Workforce Education and Development*.
[12] Difallah, D.E., Catasta, M., Demartini, G., Ipeirotis, P.G. and Cudré-Mauroux, P. 2015. The Dynamics of Micro-Task Crowdsourcing: The Case of Amazon MTurk. *Proceedings of the 24th International Conference on World Wide Web*.
[13] Hayes, G.R., Gardere, L.M., Abowd, G.D. and Truong, K.N. 2008. CareLog: A Selective Archiving Tool for Behavior Management in Schools. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*.
[14] Hedley, D., Uljarević, M., Cameron, L., Halder, S., Richdale, A. and Dissanayake, C. 2016. Employment programmes and interventions targeting adults with autism spectrum disorder: A systematic review of the literature. *Autism*.
[15] Hendricks, D. 2010. Employment and adults with autism spectrum disorders: Challenges and strategies for success. *Journal of Vocational Rehabilitation*.
[16] Hillier, A., Campbell, H., Mastriani, K., Izzo, M.V., Kool-Tucker, A.K., Cherry, L. and Beversdorf, D.Q. 2007. Two-Year Evaluation of a Vocational Support Program for Adults on the Autism Spectrum. *Career Development for Exceptional Individuals*.
[17] Hong, H., Kim, J.G., Abowd, G.D. and Arriaga, R.I. 2012. Designing a Social Network to Support the Independence of Young Adults with Autism. *Proceedings of the ACM 2012 Conference on Computer Supported Cooperative Work*.
[18] Horton, J.J. and Chilton, L.B. 2010. The Labor Economics of Paid Crowdsourcing. *Proceedings of the 11th ACM Conference on Electronic Commerce*.
[19] Howlin, P., Goode, S., Hutton, J. and Rutter, M. 2004. Adult outcome for children with autism. *Journal of Child Psychology and Psychiatry*.
[20] Keel, J.H., Mesibov, G.B. and Woods, A. V 1997. TEACCH-Supported Employment Program. *Journal of Autism and Developmental Disorders*.
[21] Kientz, J.A., Goodwin, M.S., Hayes, G.R. and Abowd, G.D. 2013. Interactive Technologies for Autism. *Synthesis Lectures on Assistive, Rehabilitative, and Health-Preserving Technologies*.
[22] Kim, J.A., Szatmari, P., Bryson, S.E., Streiner, D.L. and Wilson, F.J. 2000. The Prevalence of Anxiety and Mood Problems among Children with Autism and Asperger Syndrome. *Autism*.
[23] Leigh, J.P. and Du, J. 2015. Brief Report: Forecasting the Economic Burden of Autism in 2015 and 2025 in the United States. *Journal of Autism and Developmental Disorders*.
[24] Lindgren, S. and Doobay, A. 2011. *Evidence-based interventions for autism spectrum disorders*.
[25] Madsen, M., el Kaliouby, R., Goodwin, M. and Picard, R. 2008. Technology for Just-in-time In-situ Learning of Facial Affect for Persons Diagnosed with an Autism Spectrum Disorder. *Proceedings of the 10th International ACM SIGACCESS Conference on Computers and Accessibility*
[26] Martin, D., Hanrahan, B. V, O’Neill, J. and Gupta, N. 2014. Being a Turker. *Proceedings of the 17th ACM Conference on Computer Supported Cooperative Work and Social Computing*.
[27] Mazurek, M.O. 2013. Social media use among adults with autism spectrum disorders. *Computers in Human Behavior*.
[28] McCutchen, D. 2000. Knowledge, Processing, and Working Memory: Implications for a Theory of Writing. *Educational Psychologist*.
[29] Moffatt, K., McGrenere, J., Purves, B. and Klawe, M. 2004. The Participatory Design of a Sound and Image Enhanced Daily Planner for People with Aphasia. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*
[30] Morris, M.R., Begel, A. and Wiedermann, B. 2015. Understanding the Challenges Faced by Neurodiverse Software Engineering Employees: Towards a More Inclusive and Productive Technical Workforce. *Proceedings of the 17th International ACM SIGACCESS Conference on Computers and Accessibility*
[31] Morris, M.R., Bigham, J.P., Brewer, R., Bragg, J., Kulkarni, A., Li, J., Savage, S. and Savage, S. 2017. Subcontracting Microwork. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*
[32] Patterson, A. and Rafferty, A. 2001. Making It to Work: Towards Employment For the Young Adult with Autism. *International Journal of Language & Communication Disorders*.
[33] Picard, R.W. 2009. Future affective technology for autism and emotion communication. *Philosophical Transactions of the Royal Society B: Biological Sciences*.
[34] Ringland, K.E., Wolf, C.T., Boyd, L.E., Baldwin, M.S. and Hayes, G.R. 2016. Would You Be Mine: Appropriating Minecraft As an Assistive Technology for Youth with Autism. *Proceedings of the 18th International ACM SIGACCESS Conference on Computers and Accessibility*
[35] De Rubeis, S. et al. 2014. Synaptic, transcriptional and chromatin genes disrupted in autism. *Nature*.
[36] Seltzer, M.M., Krauss, M.W., Shattuck, P.T., Orsmond, G., Swe, A. and Lord, C. 2003. The Symptoms of Autism Spectrum Disorders in Adolescence and Adulthood. *Journal of Autism and Developmental Disorders*.
[37] Spinuzzi, C. 2005. The Methodology of Participatory Design. *Technical Communication*.
[38] Swaminathan, S., Hara, K., Bigham, J.P., Swaminathan Saiganesh, H.K. and Bigham, J.P. 2017. The Crowd Work Accessibility Problem. *Proc. of the Web for All Conference*
[39] Wakefield, J.C. 2013. DSM-5: An Overview of Changes and Controversies. *Clinical Social Work Journal*.
[40] Wobbrock, J. and Myers, B. 2006. Trackball Text Entry for People with Motor Impairments. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*
[41] Wobbrock, J.O., Kane, S.K., Gajos, K.Z., Harada, S. and Froehlich, J. 2011. Ability-Based Design: Concept, Principles and Examples. *ACM Trans. Access. Comput.*
[42] Zyskowski, K., Morris, M.R., Bigham, J.P., Gray, M.L. and Kane, S.K. 2015. Accessible Crowdwork? Understanding the Value in and Challenge of Microtask Employment for People with Disabilities. *Proc. of CSCW 2015* (2015).