

# “It’s the only thing I can trust”: Envisioning Large Language Model Use by Autistic Workers for Communication Assistance

JiWoong (Joon) Jang

jwjang@cmu.edu

Human-Computer Interaction Institute

Carnegie Mellon University

Pittsburgh, PA, USA

Patrick Carrington

pcarrington@cmu.edu

Human-Computer Interaction Institute

Carnegie Mellon University

Pittsburgh, PA, USA

Sanika Moharana

smoharan@cs.cmu.edu

Human-Computer Interaction Institute

Carnegie Mellon University

Pittsburgh, PA, USA

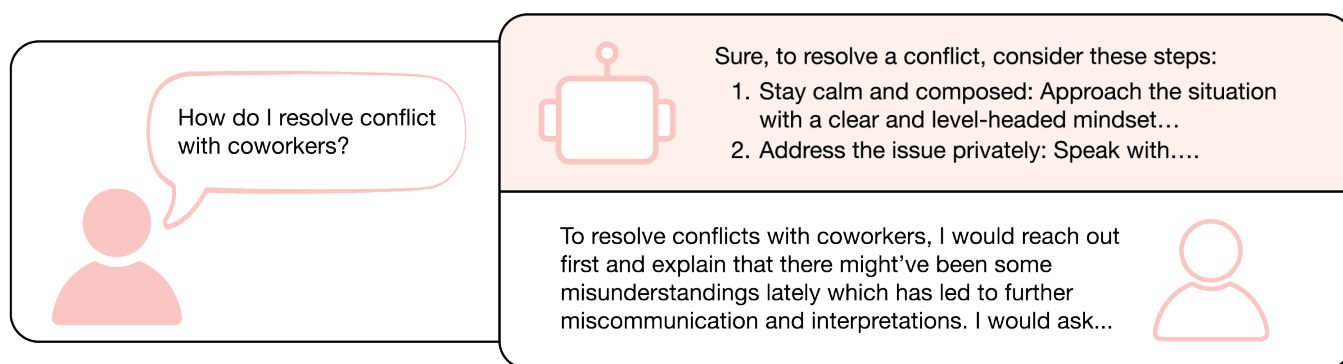
Andrew Begel

abegel@cmu.edu

Software and Societal Systems Department

Carnegie Mellon University

Pittsburgh, PA, USA



**Figure 1:** We investigate opportunities and risks from autistic workers’ use of LLMs for assistance with social communication and scenarios and had participants juxtapose LLM and human confederate-generated responses. A question from one of our participant sessions is displayed, along with corresponding LLM and human confederate replies.

## ABSTRACT

Autistic adults often experience stigma and discrimination at work, leading them to seek social communication support from coworkers, friends, and family despite emotional risks. Large language models (LLMs) are increasingly considered an alternative. In this work, we investigate the phenomenon of LLM use by autistic adults at work and explore opportunities and risks of LLMs as a source of social communication advice. We asked 11 autistic participants to present questions about their own workplace-related social difficulties to (1) a GPT-4-based chatbot and (2) a disguised human confederate. Our evaluation shows that participants strongly preferred LLM over confederate interactions. However, a coach specializing in supporting autistic job-seekers raised concerns that the LLM was dispensing questionable advice. We highlight how this divergence in participant and practitioner attitudes reflects existing schisms in

HCI on the relative privileging of end-user wants versus normative good and propose design considerations for LLMs to center autistic experiences.

## CCS CONCEPTS

• **Human-centered computing** → **Empirical studies in accessibility.**

## KEYWORDS

Autism, social communication, large language models, workplace, Artificial Intelligence, Neurodiversity

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s).

CHI '24, May 11–16, 2024, Honolulu, HI, USA

© 2024 Copyright held by the owner/author(s).

ACM ISBN 979-8-4007-0330-0/24/05.

<https://doi.org/10.1145/3613904.3642894>

## ACM Reference Format:

JiWoong (Joon) Jang, Sanika Moharana, Patrick Carrington, and Andrew Begel. 2024. “It’s the only thing I can trust”: Envisioning Large Language Model Use by Autistic Workers for Communication Assistance. In *Proceedings of the CHI Conference on Human Factors in Computing Systems (CHI '24)*, May 11–16, 2024, Honolulu, HI, USA. ACM, New York, NY, USA, 19 pages. <https://doi.org/10.1145/3613904.3642894>

## 1 INTRODUCTION

Despite wide recognition, research, and advocacy, autistic<sup>1</sup> adults continue to face critical levels of employment challenges [45, 90]. Prior work has shown that up to 85 - 90% autistic adults face unemployment and underemployment issues [71, 80], among the highest of all major disability groups in the United States [48]. When seeking employment and navigating the workplace, autistic workers report that they often face challenges such as understanding non-autistic social norms, handling limited or unclear information and expectations, discerning unspoken meaning in communication, coping with sensory overload, and managing with limited support and accommodations. Furthermore, many face difficulties navigating issues such as disclosure and advocating for roles which take advantage of their unique characteristics [10, 48, 67].

Given that autistic individuals experience challenges with social interaction in the workplace [5, 10, 67], researchers have noted the importance of having a champion or advocate who is able to mediate and assist the autistic individual in social or communication challenges at work [59]. Some research has found positive effects that specialized job coaching for autistic job-seekers and counseling with speech and language pathologists (SLPs) can have [21, 59, 74], but given the challenges and research highlighting the need for more affordable and better support [45, 67, 74, 90], it appears unlikely that autistic workers will access essential resources in most jobs. This underscores the desire for a readily accessible resource that autistic workers can turn to in moments of social difficulty – a desire which we explore whether large language models (LLMs) can support.

Our work assesses the capability of LLMs to intermediate and otherwise assist tactfully with workplace communication and related acts. This is motivated by LLMs' recent astronomic rise in popularity in adoption coupled with its apparent ability to simplify and explain social interactions. These include reports of utilization in contexts such as writing "how-tos" for social interactions, interpreting and explaining social situations, explaining humor in jokes, and proofreading workplace communications by the broader public [34]. Much of this is attributable to the fact that LLMs encode a wide range of human behavior in its training data corpus [8]. More recent chatbot incarnations (e.g. OpenAI's ChatGPT [63] and Google's Gemini / Bard [66]) have employed techniques like reinforcement learning from human feedback (RLHF) [47] to achieve an emergent capability to generate outputs with increased perceived social awareness and coherence compared to prior language models. In autism-allied and -focused spaces in social media and online forums [e.g. Facebook, X (*formerly Twitter*), TikTok, Reddit, etc.], this apparent capability has not gone unnoticed: a growing number of posts now involve users opining, speculating, and relaying their experiences after having used LLMs in social communication contexts [76]. Many explicitly note utilizing LLM-based tools like ChatGPT to understand and prepare for social situations at work [76, 84, 85], write emails and messages to supervisors and coworkers [83, 84], and understanding vague communications and instructions [83, 86] – with one thread dubbing it "a gamechanger

for people on the spectrum [86]." Given this, it can no longer be assumed that this use case is mere hypothetical.

To better envision the opportunities and concerns with autistic users utilizing LLMs for assistance in workplace communication, we aim to answer the following research questions.

- RQ 1** What communication challenges and resource availabilities (or deficits) do autistic workers experience?
- RQ 2** Do autistic workers believe that an LLM's advice could be helpful for addressing workplace communication challenges?
  - 2a** If so, why?
  - 2b** If so, how would autistic workers utilize LLMs?
- RQ 3** Can an LLM's advice be considered *good*? (and in what / whose definition of *good*?)

We conducted a within-subjects study with ( $n = 11$ ) autistic individuals where we encouraged participants to maintain a free dialogue with two chatbots, an LLM (*utilizing OpenAI's GPT-4 [64] via API with some prompt engineering*) and a human confederate in disguise. During the study, participants (1) shared with us their prior experiences with workplace communication challenges and available resources, (2) engaged in exchanges with the chatbots, and (3) rated and described their preferences for and between the LLM and confederate.

Overall, our data shows that participants showed a strong affinity to the LLM, with *nine (9) out of 11* ( $\approx 82\%$ ) expressing a desire to continue using it for communication aid and social advice. **We found the same nine (9) out of 11** ( $\approx 82\%$ ) **participants preferred LLM interactions to ones with the human confederate** through quantitative and qualitative analysis. Participants' motivations to continue using the LLM were informed by a lack of available resources at work, with many relying on friends and family despite prior experiences of emotional harm.

Participants valued the LLM for its potential to communicate in ways which ascribed to their preferences, untangle implicit neurotypical norms, and allow for the freedom to ask questions without fear of reprisal – providing a sense of control in navigating the neurotypical world. To provide grounding of the LLM's answers vis-à-vis alternative resources to autistic workers, we had a professional counselor and job coach (*LPC, NCC*) specializing in workforce readiness training for neurodivergent individuals evaluate LLM and confederate responses. While the practitioner placed value on the potential for LLMs' ease-of-access in times of need, she noted the LLM's tendency to make ungrounded assumptions and assume neurotypicality, and raised concerns about harms resulting from the LLM's misaligned advice and participant acceptance of potentially harmful recommendations.

From our results, we note a divergence in attitudes towards the LLM: **one where autistic participants express a desire to use an LLM because it appears to provide agency for independent exploration, while our practitioner urges caution due to its misleading advice.** This division symbolizes and foreshadows imminent societal concerns as LLMs are considered by autistic users [2] and are being developed explicitly for interpersonal advice [29]. Despite its novel façade, we show how the disagreement between participants and practitioner mirrors existing conflicts

<sup>1</sup>While some in the autism community prefer people-first language (PFL), others have embraced the term "autistic" as their chosen identifying label, and we have opted for that terminology as well, though we acknowledging the preference for PFL among some [43].

in HCI literature from accessibility, disability theory, and medical/social models of disability — ones involving the relative privileging of disabled experiences against normative authorities. We believe acknowledging this parallel allows for progress towards *truly assistive* aids for autistic workers which offers access to grounded advice while centering their lived experience.

This work presents the following exploratory contributions which we believe merits further consideration:

- (1) we illuminate the practice by autistic workers of obtaining social communication advice from LLMs (§ 1),
- (2) we gauge autistic participants' preferences for receiving advice from an LLM versus a human confederate and current resources in a user study (§ 4.2),
- (3) we evaluate the quality of LLM advice during a discussion with a specialized counselor and job coach (§ 4.3),
- (4) we provide discussions of potential reasons, opportunities, and concerns for future use of LLMs for autistic workers' social communication assistance (§ 5),
- (5) and we elucidate design considerations as LLMs are considered for social communications, and demonstrate how designing LLMs for social advice is fraught with entanglements and relative privileging. (§ 6)

## 2 RELATED WORKS

### 2.1 Autism in the Workplace

Autistic adults not only experience the highest level of unemployment amongst all major disability groups in the United States, but they also tend to be the least considered in any company diversity or disability efforts, reinforcing the large employment gap for autistic adults [33]. When seeking employment and navigating the workplace, autistic workers report that they often face challenges such as overwhelming sensory processing limitations, and accessing accommodations [71, 90]. Apart from these, they face greater invisible social difficulties including understanding non-autistic social norms, ambiguous expectations, discerning unspoken meaning in communication [59] that result in negative experiences [50, 78]. These contrasting communication practices and preferences between autistic and non-autistic people pose a mismatch across social interaction as exemplified by Milton's Double Empathy perspective [56].

Research has highlighted the significance of autistic individuals having a champion or advocate in their corner to assist with workplace communication challenges [61], but they also might turn to family, friends, and peers for daily living activities and advice, or find other forms of social support networks online that help meet everyday socioemotional needs [16, 25]. However, these personal resources do not always possess the aptitude to provide guidance for workplace dilemmas.

Within the workplace, the landscape of resources for autistic people in mitigating these difficulties can include job coaches, vocational rehabilitation specialists, occupational therapists or licensed professional counselors (LPC's) [21]. These experts can provide personalized support for autistic individuals across their employment, career pathways, and professional development journeys, but are largely inaccessible and not available at scale for autistic

adults [30, 91]. There are also community organizations, government agencies, and nonprofits that offer employment readiness programs and support services [24, 60]. Many workplaces are integrating wider autism at work initiatives and efforts within human resource departments to advance assistance [41]. However access and availability of workplace accommodations vary across domains, industries of work, and their organizational policies. With the number of adults with autism seeking support increasing, many are unable find resources, services, and programs through the aforementioned avenues [13].

Considering the discrimination due to communication differences that autistic individuals experience, most development of technological aids and tools are directed towards social engagement, educational goals, and learning experiences [92]. Many online communication platforms, internet forums, and social media have been beneficial for autistic people in cultivating connectedness, social relationships, and seeking support in a controllable manner [14, 35], especially through written and text-based mediums [39]. Emerging technologies that have surfaced hyperfocus on autistic children and support for parents in promoting social communication via voice assistants, smart glasses, AAC devices, and other speech-generated outputs [44, 72, 73]. With popular investment of early interventions for autistic children [38, 97], there is an increasing amount of negligence towards autistic adults and a lack of tools for supporting their social communication needs [18]. In the context of the workplace, interactive communication technologies aimed to help adults in interviewing, the job search, and pre-emptive training process can be occasionally found, [36], but tools for longitudinal support and retention in the workplace are scarce [6].

Another critical element in finding support is dependent on disclosure. Disclosure of one's autism can be an increasingly vulnerable and stigmatizing experience even amongst immediate friends and family [22, 81, 82]. Disclosure in the workplace poses further potential risks to job security, unwarranted biases and microaggressions, and threats to acceptance [48]. Within workplaces, the situation may vary — larger and more structured workplaces have designated HR professionals to support employees in requesting and receiving accommodations, while others may mandate speaking directly to a supervisor or manager. This establishes more power differentials and fear of transparency, placing a greater burden on autistic employees to navigate on their own. Considering these disclosure challenges [49], coupled with the plethora of research which speaks to the need for improved affordable support structures for addressing social difficulties, it is unlikely that many autistic workers would have access to the kinds of resources they need.

### 2.2 HCI and Autism

Within HCI, designing for neurodiversity and technology for supporting autistic individuals has been explored in most popular applications through web search and navigation, virtual reality, software development, visual tools, and especially computer-mediated communication tools such as social network and media platforms, video calling, texting and messaging [14, 27, 46, 79, 98]. There has been research devoted to conducting participatory design work with autistic individuals and assistive technology, identifying the challenges that inclusive and universal design methods face when

designing for autistic people and ways of better collaborating with neurodivergent adults. [88]. However, significant amounts of HCI research relating to autism have focused on autistic children, who have different communication patterns and needs than adults.

Within the workplace context, Annabi et al. have characterized autism employment across experiences for women, software developers and other IT professionals while creating frameworks and context through which adjustments, accommodations, and support for these autistic workers can be explored to create more inclusive learning and mitigating such barriers [3, 4, 69]. Reviews of accessibility research in HCI [51] and particularly neurodiversity research, centers around the mismatches of neurodivergent users using generalized technology or how assistive technology can drive behavior change for these, but few focus on consequences of such interventions.

### 2.3 Autism, Generative AI, and LLMs

In the rapidly evolving landscape of generative AI and LLMs, HCI research contemplates the possibilities of its application towards various domains of use, along with developing guidelines [93] and exploring its boundaries. The implications for capability, agency, and control are questioned as the quality of generative AI outcomes begin to parallel or at times surpass human strengths [57]. While some researchers imagine its contributions as a design space, material, and avenue of prototyping for human-centered computing systems [68], a lot of attention has been devoted to how its changing the landscape of learning and educational systems [42]. Others have suggested its potential for supporting communication for people with disabilities, such as enhancing experiences for AAC users [87], writing AI-assisted emails for adults with dyslexia [28] and even speech-language pathologists in delivering therapy to autistic individuals in clinical settings [20].

However, researchers also caution towards malicious uses, negative impacts, and reliability of LLM-generated content over its low cost and high speed [26, 31]. Its efficiency is contrasted with the ethical boundaries of human augmentation or automation [15] for future applications. Gadiraju et al. provide disability-centered perspectives on LLMs through discussions with people with various disabilities to formulate more disability-positive data representation strategies for diminishing societal harms and biases perpetuated by such neural language model-based dialog systems [23]. Despite these modern advances, it remains the case that the risks highlighted by Bender et al. [9], of exposure to encoded biases and discriminatory viewpoints, lack of grounded communication, and overreliance on automated systems likely still apply with current and future human-LLM interactions, potentially exposing autistic users to said harms [94].

### 2.4 Viewpoints on Technology's Role, Design, and Relationship with Disability

Because the subject matter of this work involves a technology's role in managing a disabled person's social representation, we introduce viewpoints which speak to worldviews on disability, and subsequently, technology's role, particularly when its harms are taken into account. We do so to acknowledge the discourse around changing definitions and interpretations of disability, which have

prompted many to reassess how research includes or excludes people with disabilities' representation and how as it maps onto technology design and development.

*The Medical Model of disability*, as the most mainstream perspective, views disability as something an individual tackles for themselves, and generally advocates for utilizing existing medicalized structures, such as therapy, or counseling [12]. From this perspective, some help, despite being coupled with a loss of individual agency, may be seen as better than no help at all, especially if the benefits to said technology aligns with the model's focus on management and mitigation of disability-related challenges [53]. The model might also lean towards the goal of assistive technology being designed for an organizational or societal good, as it often promotes strategies that help individuals conform to existing structures and norms.

In contrast, the *Social Model of disability* sees disability as a societal issue caused by barriers and attitudes within society and in general would reject the premise of relying on technological help for human-human social assistance — arguing for a societal or organizational-encompassing solution to address unmet needs of disabled individuals [62, 75].

Relatedly, *Standpoint Theory*, advocated by scholars like Lucy Suchman and Shaowen Bardzell and originating from feminist theory, remains influential in current thinking on technology design processes. It highlights the significance of individual perspectives and experiences, emphasizing that the primary objective of assistive technology is to advance the interests of the individual [7, 77].

*Disability Studies and Critical Disability Studies*, drawing from works like those of Mankoff et al., Hamraie, and Williams et al., go further, with the goal for technology to challenge existing power structures and narratives about disability and advocate for amplifying and enabling the voices and experiences of disabled individuals [32, 52, 95]. These perspectives serve as ones which most directly challenge the medical model of disability's view on technology as being permissive or enforcing the status quo.

Each of these perspectives relate to differing goals and resulting tradeoffs they posit that assistive technology ought to exhibit. We explore the implications of these perspectives as we investigate autistic workers' potential use for LLMs in social and communication assistance.

## 3 METHODOLOGY

In the study, autistic workers were asked to share prior workplace communication challenges and resources, and to interact with and rate two chatbots: (1) *Paprika*, a chatbot which utilized OpenAI's GPT-4 [64] API (*with some modifications described in § 3.1.4*) and (2) *Pepper*, a human confederate. A study design with a human confederate was chosen as a comparative baseline for evaluating the LLM as it represents a close analog for a readily available resource that provides human-like advice, akin to an anonymous help hotline (in chat form), Reddit comments, the social network Blind [1], as well as interaction behavior which were described in social media posts and threads. To ensure participants would feel comfortable providing candid feedback without worrying about potentially offending a real person, we communicated that both *Paprika* and *Pepper* were automated agents. After the conclusion of the study period, we

consulted with a professional counselor and job coach (*LPC, NCC*) specializing in workforce readiness training for neurodivergent individuals for their opinions to reveal additional context into the efficacy and safety of LLM-generated advice.

### 3.1 Study Protocol

The study, whose design was approved by an Institutional Review Board (IRB), was conducted remotely, and was designed to take 90 minutes across three parts for each participant. Data was recorded via video conference software recording, and collected video and audio data from participants, as was text inputted to the Discord space used for chatbot interactions and surveys (§ 3.3). Each participant encountered the following three stages:

- (1) a 10-minute semi-structured interview (§ 3.1.1),
- (2) a 65-minute session interacting with and evaluating two chatbots (§ 3.1.2) which included,
  - a) a 5-minute onboarding to familiarize the participant with Discord and interacting with *Paprika* (LLM) and *Pepper* (human confederate),
  - b) four (4) sessions of an (up to) 10-minute interactions with a chatbot and a rating survey,
- (3) and a 15-minute follow-up semi-structured interview (§ 3.1.3).

After the study, participants were disclosed of the study’s purpose, and that *Pepper* was a human confederate, and further follow-up questions were addressed if raised.

**3.1.1 Assessing Prior Experiences.** In an initial 10 minute semi-structured interview, participants were asked about prior issues they had encountered with workplace communication, either in person or online, as well as the context surrounding any of these challenges. Participants were also asked about types of assistance they had sought or considered (in or outside of work) when faced with these issues. Separately, participants were asked whether they had used chatbots or LLMs in the past.

**3.1.2 Interaction and Evaluation of Chatbots.** Participants were navigated through the process of opening the Discord client on their device and connecting to the server used for the experiment. There were then introduced to the different rooms: one for testing/trialling (connected to an LLM) and four for interactions with the two chatbots (two each per chatbot), with ordering done in an alternating and counterbalanced fashion (including between participants). Participants were introduced to a scenario where they would attempt to ask various chatbots for advice, and that they were to consider the chatbot to be disassociated with any workplace software, and thus the agents would not have context of who they were, and also would not report data back to their workplace.

**Onboarding – Test Interaction Room.** In the first room, participants were encouraged to ask the chatbot any kind of question (e.g. “What’s your favorite recipe for chocolate chip cookies?”) to acclimate participants to the idea that the chatbots in this study could be more conversant than those which could be found in commercial applications, such as customer service chatbots. Most participants (P1 - P3, P6, P7, P9 - 11) opted to ask workplace communication related questions (“What are some strengths of neurodivergent people in technology?” - P8) while some (2) chose to ask philosophical

questions to challenge the system (“What is the meaning of life?” - P4, P5).

**Interactions with *Paprika* and *Pepper*.** Before each individual interaction, participants were informed that each room was connected to either the *Paprika* (LLM) or *Pepper* (human confederate), and participants were encouraged to discuss their prior workplace communication-related questions with the chatbots. Generic questions were prepared, but all participants opted to ask their own questions. After each of the four (4) chatbot interactions, participants completed a short survey (described in § 3.3). Participants interacted with *Paprika* and *Pepper* consecutively and in alternating fashion, with a 10-minute limit per interaction.

**3.1.3 Collecting Overall Impressions.** After completing the interactions and evaluations of both chatbots, participants engaged in a follow-up semi-structured interview. During this interview, they were asked to reflect on their experiences with both *Paprika* and *Pepper*, provide comparisons between the two, and discuss any preferences or suggestions for improvements. Additionally, participants were asked to share their thoughts on the potential utility of chatbots in assisting with workplace communication issues and any concerns they might have regarding the use of such technology. Upon conclusion of the follow-up interview, participants were debriefed about the true nature of *Pepper* being a human confederate and of the purpose of the study. They were given the opportunity to ask any questions or share any additional thoughts about the study design and their experiences.

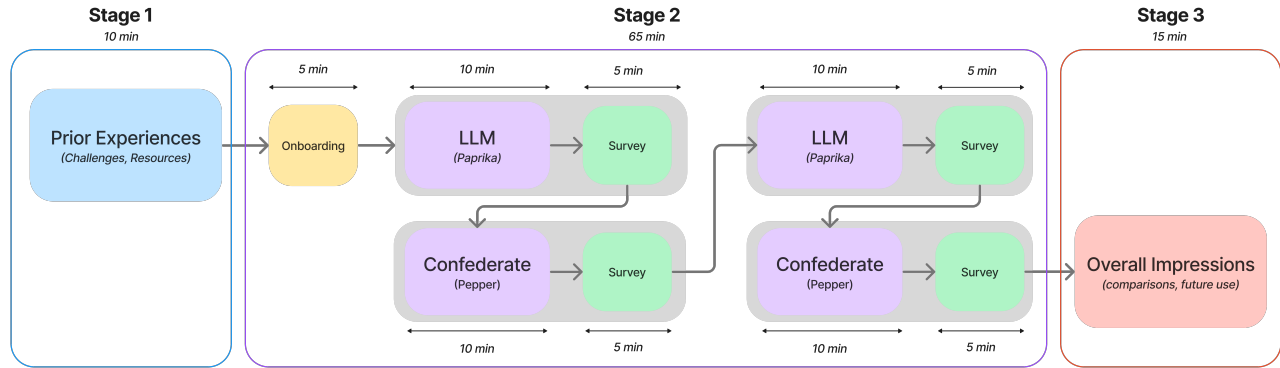
**3.1.4 *Paprika*: the Large Language Model Chatbot.** *Paprika* was developed using OpenAI’s GPT-4 [64] API, with it prompted with the following before each 10-minute interaction:

- (1) *You are a helpful assistant named Paprika. Provide clear and thorough answers but be concise.*
- (2) *Use a more conversational but still workplace appropriate style. Make sure your answers are short, make sure your responses are around two paragraphs.*
- (3) *Also, if I am not asking a question that is workplace-communication related, let me know that I am off-topic and steer the conversation back on-topic to workplace communication. Do not attempt to answer the question if it is off-topic.*

to ensure some parity with the human confederate *Pepper* in response length and writing style. (Note that the prompt is demarcated in sections for reading clarity – in the study, the sentences of the prompt were formatted as a single prompt, in a single paragraph.)

**Fine-Tuning *Paprika*’s Prompt and Behavior.** Initially, the study originally considered only text from the beginning of the final prompt (e.g. (1)). After some testing with our confederate, the second phrase (2) was included to help ensure responses were comparable to response styles that the human confederate (*Pepper*) provided. Finally, the last phrase (3) was added to have *Paprika* (LLM) only respond to questions germane to workplace communication, as was instructed to our confederate (*Pepper*).

For the first half of studies (P1 through P6) responses from *Paprika* were posted after receiving a response from the API and undergoing human review by the interviewing researcher for safety.



**Figure 2: A diagram outlining the study protocol used. The intervention ordering was counterbalanced between participants and participants engaged with a different chatbot each session.**

In the later trials (the last five trials), after noting some participants were sensitive to the response time difference between *Paprika* and the confederate [mean delay for first 6 participants: *Paprika* = 165 seconds, *Pepper* = 189 sec], output from *Paprika* was delayed to match or exceed the delay observed from *Pepper* (the confederate) from previous trials.

Similarly, *Paprika*'s verbosity was manually limited to outputs which were within 30 words of the confederate's average reply length from previous participants [mean verbosity for first 6 participants: *Paprika* = 205 words, *Pepper* = 153 words]. This was achieved by continuously re-sending API queries until a response of desirable length was received. (Note that this only occurred once in all interactions with the last five participants.)

**3.1.5 Pepper: the Human Confederate.** The human confederate, who posed as the chatbot *Pepper*, is a graduate student with three (3) years of prior work experience. In preparation for this study they consulted with a practitioner (§ 3.4) about best practices as a support for a neurodivergent coworker. During the study, the confederate did not have access to content from conversations with *Paprika*, and was tasked to answer questions from participants as they were entered into the Discord channel.

## 3.2 Participant Details

We recruited eleven (11) participants, whose ages ranged from 22 to 50 (mean = 30, SD = 8.7), and with a wide variance of educational and work experience from 0.5 to 28 years (mean = 9, SD = 8.1). We report demographics and work experiences of our participants in Table 1. Participants were recruited by posts to general-purpose and special-purpose (e.g. autism advocacy related) email lists, as well as with direct outreach. Promotional and recruiting materials for the study referenced the potential of interacting with and rating chatbots based on LLMs and which may be similar to ChatGPT. Participants were informed that their responses to surveys, chatbots, as well as video and audio from the interview may be used for data analysis. After the study, participants were compensated \$20 in gift card credit for their time.

## 3.3 Data Collection

Overall, data collected from this study included (1) transcripts from recordings of the semi-structured interviews, (2) text from the chatbot interactions, and (3) participant responses to the post-interaction surveys

**3.3.1 Post-Interaction Survey Design.** The post-interaction survey given to participants after every (up to) 10-minute interaction included questions on a 7-point Likert scale assessing participants' perceptions of the chatbots' *utility*, *understanding*, *likelihood of use*, and *dependability*. Details about the questions and the anchors to the Likert scales are provided in Table 2.

As part of the survey, we also asked participants the following questions, which they could provide long-form written answers.

- (1) If you could change this chatbot's behavior, or give it feedback, what would you tell it?
- (2) What concerns, if any, do you have about your interaction with this chatbot?
- (3) How does this chatbot's response differ from what you may get from a coworker / friend / supervisor if you asked them for advice? (and tell us what/who you are comparing against)

**3.3.2 Quantitative - Overall.** When participants were asked which chatbot they preferred overall, **9 out of 11** ( $\approx 82\%$ ) expressed a preference towards the LLM (*Paprika*) over the confederate. We conducted a two-tailed Mann-Whitney U / Wilcoxon's signed-ranked test on the participants' Likert scale ratings for *Paprika* and *Pepper* to determine if there was a statistically significant difference in participant ratings between the two. Given that *Paprika*'s (LLM) response time and verbosity were altered for the latter five participants, we conducted additional two-tailed Mann-Whitney U / Wilcoxon's signed-ranked tests on the first six and latter five participants separately to ascertain whether statistically significant differences were observed in either or both groups.

P#	Age	Gender	Ed. Level	Work Exp.	Types of Work
P1	22	F	bachelor's degree,	3	food service, library management
P2	20	F	some college credit	2	veterinarian hospital assistant, food service
P3	24	F	bachelor's degree	4	food service, mental health counselling intern, dorm residence advisor
P4	33	F	doctorate degree	9	Ph.D. student, lab manager, postdoc
P5	33	M	bachelor's degree	5	circus actor, tech company analyst / manager, factory machining
P6	26	F	master's degree	2	data science support engineer
P7	34	M	master's degree	16	personal business, research staff, Ph.D. student, contractor work (tech), corporate tech developer
P8	30	F	some college credit	6	call center customer service
P9	38	F	bachelor's degree	17	project management, dog kennel
P10	50	F	professional degree	28	investment banker, university staff, tax consulting, medical consulting
P11	24	F	master's degree	5	dorm residence advisor, field officer for NGO, research assistant, psychology instructor

Table 1: Participant Demographic Data

Rated Quality	Question Text	Scale Extrema
Utility	How would you rate the helpfulness of responses from the chatbot?	Not at all helpful (1) / Extremely helpful (7)
Understanding	How would you rate how well the chatbot understood you?	Did not understand at all (1) / Understood extremely well (7)
Intent to use	Would you use this chatbot for your personal use in the future?	Definitely no (1) / Definitely yes (7)
Dependability	How dependable do you find this chatbot?	Not dependable at all (1) / Extremely dependable (7)

Table 2: Likert scale questions administered after short chatbot interactions.

3.3.3 *Qualitative - Overall.* Long-form answers to the survey questions, as well as transcribed participant dialogue from the semi-structured interviews in (§ 3.1.1) and (§ 3.1.3) was utilized for thematic analysis [11] – to identify patterns and themes in the qualitative data. The analysis began with two researchers jointly reviewing three (3) transcripts and corresponding survey responses and discussing the development of a codebook. Once the codebook was agreed upon, one of the two researchers coded the remaining transcripts. After completing this coding, both researchers met again to review the coded data, discuss any discrepancies, and finalize the identified themes and patterns. Our thematic analysis of the participant survey responses and interview transcripts yielded a codebook with 29 codes, which we report in Table ?? in the Appendix.

### 3.4 Practitioner Evaluation

After completing all 11 participant sessions, we sought additional grounding and points of comparison for LLM-generated advice versus other known resources, and we invited a professional counselor and job coach (*LPC, NCC*) who specializes in workforce readiness training for neurodivergent individuals to review the responses from both *Paprika* and *Pepper* and provide expert validation [54] of LLM-responses from the study.

Over a 150 minute session, the practitioner reviewed chat transcripts from the participant interactions for both the LLM and the confederate and gave open-ended feedback as to the quality of the responses and comparisons to advice she would give in her practice. The practitioner conducted back and forth dialogue with the researcher and was encouraged to think-aloud while she reviewed the chat transcripts, and engaged in an unstructured exploratory problem-centered interview [96] about the practical effects of more widespread access and use of LLMs and of access to advice of this form.

## 4 RESULTS

We address our research questions with findings from our quantitative, qualitative, and expert interview data.

### 4.1 Participant Experiences in Workplace Environments

4.1.1 *Communication Norms and Challenges.* Participants reported facing a large variety of communication challenges in the workplace. The most frequently cited challenge involved navigating ambiguity. All (11) participants shared their difficulties with gaining what they perceive the necessary clarity to proceed on a work objective or communication, and the resulting judgemental reactions from



coworkers and supervisors for asking for more clarity, similar to what P1 relates here:

**P1:** *[When I ask for specifics,] it kind of comes off like [...] I need more hand-holding when it comes to these tasks, but I don't. I just need better descriptions [...] and it comes off as if I don't know what I'm doing.*

Many participants (8 | P1 - P4, P6 - P8, P10) related to feeling self-conscious about their need for additional clarification beyond what neurotypical coworkers would ordinarily request. In addition, participants also spoke to difficulties with ascertaining communication norms, much of which were implicit. P7, for example, recounted being confused by “the different [expectations] between being in a meeting room versus chatting with someone at their desk.” These same participants (8 | P1 - P4, P6 - P8, P10) bemoaned the unwritten rules in workplace communication, much of which they learned by running afoul.

These challenges created substantial difficulties for participants which often extended beyond working hours – people shared that they felt exhaustion (P3, P5, P7), anger (P2, P8, P9), anxiety (P2, P3, P11), frustration (P10), and battled with the urge to overthink (7 | P2, P3, P4, P6, P7, P9, P11). P10 described how these challenges haunted her throughout her adult life:

**P10:** *In the end, it feels like the fault is always me. The common denominator is me. It's like, [...] what's wrong with her? What's wrong with her? I think about that all the time.*

**4.1.2 Job Environment.** Our participants as a group had experiences with a wide variety of job settings (Table 1). We observed that participants asked questions about spoken and written communications (whether synchronous or not) in roughly equal proportion, suggesting interest in assistance in both modalities. Participants who associated their job environment factoring into their communication challenges relayed a multitude of different points of frustration including bright overhead lights (P3), a requirement to attend a large professional conference (P9), and unpredictable participation in group interaction and discussions (4 | P2, P3, P4, P8). Meanwhile, participants who expressed satisfaction with their current or past job environments (4 | P2, P3, P7, P10) all mentioned flexible scheduling and choice of work environment, which included communication channels with a supervisor or Human Resources (HR) staff who were open to accommodating participants' needs.

**4.1.3 Support and Resources.** Participants varied widely in the supports they utilized and their success in addressing their challenges when doing so. Some (P1, P3, P4) like P3, found supervisors and coworkers who would be willing to consistently answer clarification questions and to provide accommodations as needed. Several (P2, P9), like P2 found that “[supervisors] said they would [...] but then they would brush it off and not come to help.” Meanwhile, others (4 | P5, P8, P9, P10) attempted to utilize workplace-based resources, such as a Human Resources (HR) office, but all of these participants expressed frustrations with getting accommodations, and some (4 | P2, P8, P9, P10) even reported being ignored or outright denied help.

To address their needs, participants reported attempting to utilize personal resources like friends, family, and mentors for communication advice. P7 shared that he found asking for assistance from friends and mentors to be particularly fulfilling, as it allowed him “to show [his] friends that [he] respect[s] their opinion and [...] build relationships with other people.” Not everyone shared this view. Several participants (5 | P1, P2, P4, P8, P11) had privacy-related qualms or shared stories of experiencing emotional distress after sharing communication struggles connected with their autism. P1 described her mother's reaction as “50/50 dangerous” due to her not understanding common difficulties faced by autistic individuals, while P8 described her husband's advice and reactions as “pretty hurtful” and refuses to consult with him with workplace communication challenges. Still others (P6, P7, P9) shared that they attempted to understand how to address challenges by looking at online self-help articles and books on workplace communication. P7 (as well as P9) stated that he liked to browse Reddit threads to see users' exchange of ideas and come to his own conclusion.

Perhaps what was most concerning was that many (5 | P1, P2, P8, P9, P10) of our participants communicated that they did not consider themselves to have any resources to rely on for workplace communication assistance. Some of this was driven by participants' desire for their source of advice to understand autistic perspectives, as P1 explains:

**P1:** *When people talk about communication errors, I think they come from a very able-bodied, neurotypical sense. [...] I feel like (most assistance) never addresses my specific needs or the needs of my community.*

Most crucially, participants connected their lack of resources to their knowledge gaps on difficulties on finding out how to approach their challenges. As P9 shared:

**P9:** *I've asked other people for advice, [...] observed people throughout my life, [...] read specific books, [...] watched videos, [...] I've explored these things. I think I just haven't found the right approach, [...] and I haven't found the best information to help with that.*

## 4.2 Did Participants Find the LLM Helpful?

Prior to the study, five (5 | P4, P5, P6, P7, P11) out of 11 participants reported having had prior awareness of large language models in some context. Three (3 | P4, P6, P7) had previously attempted to utilize ChatGPT or similar models in workplace communication contexts. With this in mind, we report our participant's responses regarding their perception of the helpfulness of the interactions from the LLM.

### 4.2.1 In Comparison with the Human Confederate and Current Resources.

*Quantitative Findings from Comparisons of the LLM and Confederate.* At the conclusion of the study, when participants were asked which chatbot they preferred overall, **9 out of 11** ( $\approx 82\%$ ) (P1 - P4, P6, P8 - P11) preferred their interaction with the LLM (*Paprika*) over the confederate (*Pepper*).

With participant ratings from post-interaction surveys, we report that participants rated *Paprika* (LLM) more highly than *Pepper* (human confederate) for all four questions at the  $p < 0.05$  level,



	All Sessions			No Adjustment (P1 - P6)			Verbosity & Latency Matched (P7 - P11)		
	LLM	Conf.	p-val	LLM	Conf.	p-val	LLM	Conf.	p-val
<b>Utility</b>	5.75 (0.85)	4.2 (1.44)	0.00024**	5.82 (0.98)	4.82 (0.75)	0.014*	5.67 (0.71)	3.44 (1.74)	0.004**
<b>Understanding</b>	6.05 (1.15)	4.75 (1.48)	0.0037**	6.00 (1.34)	5.18 (1.25)	0.041*	6.11 (0.93)	4.22 (1.64)	0.0087**
<b>Intent to Use</b>	6.00 (1.26)	4.35 (1.26)	0.0024**	5.64 (1.36)	4.36 (1.21)	0.031*	6.44 (1.26)	4.33 (1.26)	0.043*
<b>Dependability</b>	5.55 (1.10)	4.4 (1.64)	0.013*	5.55 (0.93)	5.09 (0.74)	0.079	5.56 (0.88)	3.56 (1.94)	0.017*

**Table 3: Mean ratings on the tested four attributes in the post-interaction survey are displayed, with standard deviation in parentheses. All participant ratings, ratings from participants interacting with the LLM with and without adjustment are provided separately. Our p-values were calculated using a two-tailed Mann-Whitney U / Wilcoxon’s signed-ranked test on participants’ Likert scale ratings.**

**:= significance at the  $p = 0.05$  level, \*\* := significance at the  $p = 0.01$  level**

which was true even in cases where each of the counterbalanced interaction order of chatbots were tested individually. Additionally, though the LLM was modified after six (6) participants to even its verbosity and latency with the confederate (§ 3.1.4), we nonetheless observed a statistically significant preference toward the LLM in both conditions in almost all cases and questions. Specifics are provided in Table 3.

As a snapshot of a moment in time (given the rapid development of today’s state-of-the-art models), we believe these quantitative results help demonstrate the possibility that some autistic workers may in many cases (or even outright) prefer asking social advice from an LLM currently, and to lend credence to the idea that this preference is quantifiable.

*Qualitative Findings of Nuanced Reasons for Preferences.* Participants (9 | P1 - P4, P6, P8 - P11) commented that they liked the LLMs’ ability to provide well-structured answers, many with step by step instructions or breakdown of the problem, with P3 noting how LLM responses were “far more visual[ly structured] than what I got from [the confederate].” This led to the perception that the LLM was giving more clear advice than the confederate. Some (5 | P1 - P3, P8, P10) attributed this clarity to the structuring of text output, but others (4 | P5, P6, P7, P9) noted that the LLM output was slightly more verbose. Some (3 | P5, P6, P8) found the sometimes lengthier responses from the LLM to be overwhelming at times, but of those three, two (P6, P8) mentioned that they’d “rather have the detail than not” (P8).

In addition, most (7 | P1 - P4, P6, P8, P10) participants responded and commented that they felt that the LLM understood their requests better than the human confederate. However, some (P5, P7) participants reacted strongly in the opposite manner. This division was attributable to the LLM’s tendency to assume the participant’s context and provide immediate response while the human confederate often asked follow-up questions to gather more information. Assumptions like this on the part of the LLM was a cause of concern for some participants (4 | P5, P7, P8, P11), in that it gave them less confidence in the applicability of its advice. Yet the confederate’s behavior (and the subsequent increased interactions required to reach an answer) served as a disincentive and friction for others (7 | P1, P2, P3, P4, P8, P9, P11).

With respect to comparisons against current resources, participants’ comments on the LLM interactions largely echoed those

from the comparison against the confederate. For participants who wanted to continue using it (9 | P1 - P4, P6, P8 - P11), the LLM was seen as a possible tool for independent exploration and “neurotypical worldbuilding” (P9). These participants liked the idea of how LLMs could allow them to prepare and strategize for particularly stressful communication related scenarios without reprisal.

*4.2.2 Attributes of the LLM that Participants Liked.* With respect to specific attributes, participants cited the following, in order of most number of participants to least:

- (1) *Improved Formatting* (9 participants | P1 - P4, P5, P7 - P10): The LLMs’ tendency to organize and present content into numbered or bulleted lists and providing (bordering on) verbose descriptions led to participants’ perceptions that the LLM understood them and their query better, as well as gave more comprehensible advice.
- (2) *Expedience* (7 | P1 - P3, P6, P8 - P11): For similar reasons, many participants commented that they preferred the LLM over the human confederate because of how quickly it could respond to questions relative to both the human confederate and existing sources for social guidance.
- (3) *Privacy* (6 | P1 - P4, P7, P8): Many of our participants remarked on the importance of having a source in which they could ask social guidance questions in confidence. Participants affirmed the importance of privacy for future iterations of chatbots where social advice for work could be sought.
- (4) *Open Mindedness* (4 | P3, P6, P8, P9): While we discouraged participants from asking questions which were not directly germane to workplace communication and created mitigations for this possibility, some (P3, P6, P8, P9) nonetheless succeeded in asking questions related to other workplace activities, such as generating good presentation topics for an upcoming meeting. These participants expressed pleasant surprise that the agent promptly replied and supported detouring into different topics without making judgemental remarks.
- (5) *Conversational Tone* (4 | P1, P4, P8, P10): Some (P1, P8) were drawn to the LLM’s tendency to include polite expressions,

such as “Certainly!” or “Of course!” Relating to this, P1 remarked: “It’s awfully earnest, and I don’t get that a lot.” Others (P4, P10) echoed this sentiment and how the LLM’s conversational tone made them feel as their questions were valid, even welcomed.

- (6) *Convenience/Availability* (3 | P1, P2, P11): Some touted the potential benefits of having some source of assistance readily available – “I would go to *Paprika* all the time if I could get it on my phone through Discord, just in case, you know?” (P1)
- (7) *Affordability* (3 | P1, P4, P9): A subset of participants reported utilizing or considering seeking a practitioner such as a job coach with difficulties encountered with workplace communication. As such, when it was mentioned that agents similar to the LLM used could be accessed for free (e.g. *ChatGPT*), or for relatively minimal cost (*USD \$0.03 for a roughly 250-word back-and-forth exchange, as of Aug 2023 via the GPT-4 API*), these participants expressed excitement at the potential for relatively affordable access.

**4.2.3 Envisioning the Role of an LLM Chatbot.** At the conclusion of the study, participants were asked to describe if the LLM were available for use, how it would fit within their array of support resources. Seven (7 | P2, P3, P4, P8, P9, P10, P11) participants responded that they would consider the LLM a primary, go-to resource for mediating workplace communications. P9 (along with P4, P8, and P10) described that she would utilize it as she didn’t have access to other resources (both personal and work-related):

**P9:** *Well, I think honestly with my workplace, it would probably be... it’s the only thing I can trust because, not every company or business is inclusive.*

Others like P1 and P6 considered LLMs as “a place to get [a] second opinion” (P6). Two (2 | P1, P6) described that they would consider using the LLM on occasion, while two (2 | P5, P7) expressed hesitation for using it for much of the specific challenges they were facing. Both participants explained that their reluctance resulted from a desire to make sure “[their] voice is actually articulated” (P5) in situations of resolving interpersonal conflict.

### 4.3 A Practitioner’s Perspective on LLM Advice

Our expert review [96] with a counselor (LPC, NCC), who currently directs a job-training assistance program for autistic students at a university, yielded several points which adds nuance to interpreting the quality of the large language model’s feedback for an autistic audience. We present significant observations present throughout her feedback and corresponding commentary.

**4.3.1 Overall, Similarity to Advice in Practice but Often Misleading.** Frequently, the counselor would remark that the LLMs would provide advice similar to or exactly like ones she would provide in practice.

On reading a response from the LLM about the need to document (e.g. preserve written communication) efforts towards conflict resolution with a supervisor:

**Counselor:** *This is great, actually, perfect. [...] The documentation part, I think, is important to mention.*

Overall, she rated the LLM as doing well at dispensing general-purpose advice, but found that many autistic individuals may have

difficulty with adapting the advice to their personal context which could lead to compounding ramifications, which we explain below.

**4.3.2 Advice Rooted in Neurotypicality.** The counselor noted some of the LLM responses in which participants asked about how to approach social situations included neurotypical-normative approaches to solving challenges. Some, she notes, may lead to additional confusion or misunderstanding by the autistic user.

**Counselor:** *[The LLM is] saying, start by being approachable and start a casual conversation. So like, what does that mean? Neurotypical culture would be like, make appropriate eye contact, laugh at people’s jokes, things like that, or just have open body language. But for somebody who’s autistic, eye contact is not [easy]. But also, even knowing what approachable means can be hard.*

Likewise, for instances when the LLM recommended that the participant “understand the other party’s perspective,” the counselor noted that this advice may not be applicable for many.

**Counselor:** *I think the application of it might be a challenge for some. It’s like, how do you execute, like actually going about doing it – might be hard.[...] If your boss or supervisor is not aware of neurodiversity and they’re not very open and understanding – I think that’s even more of a challenge to understand their [the supervisor’s] perspective and work with them.*

In addition, she explains that that certain advice, while commonly heard in neurotypical contexts, may be situationally harmful to the autistic individual and require additional clarification to be actionable.

**Counselor:** *[The LLM says] embrac[e your] unique qualities and focus on building genuine connections with your coworkers. I can see a lot of people who are autistic or just like are socially different, their unique qualities are such that, they can’t be their unique self without it causing problems – that’s been their experience in the past.*

**Researcher:** *So if they follow this advice, that could lead to more problems?*

**Counselor:** *I think for people that are having more significant challenges, I think [this advice] would not necessarily be enough information and might be going against what we’d advise and have experienced.*

**4.3.3 Blurring of Employer-centric and User-centric Advice.** Given an interaction where the large language model recommended that P2 (a veterinarian’s assistant) should “always keep the patient’s well-being and satisfaction as your top priority” when handling a patient confrontation, the counselor noted that this recommendation could pressure the participant into pushing past their work-life boundaries and cause them to overexert themselves.

**4.3.4 Tendency for Optimism.** The counselor observed that when participants asked for advice with planning a challenging social encounter, such as a meeting with a supervisor, or resolving a conflict with a coworker, the LLM tended to assume the best-case scenario would occur. This could result in a situation where the

autistic individual is unprepared for when an interaction goes off-script from what the LLM predicts, particularly if the conversation partner does not agree with the autistic individual.

In an exchange where the participant asked about how to ask for accommodations from their supervisor, she noted:

**Counselor:** *I also think that it would get them in trouble when it doesn't go smoothly. Because what do you do if it doesn't go well? If they know to ask those questions, then they might (be able to) mitigate that. But I think for people that are like, I don't know how to talk to my supervisor about accommodations or something, and they [the LLM's advice] and they run with it, their supervisor [refuses] [...] I think that's where they might get into trouble.*

**4.3.5 Importance of In-Situ Guidance and Help.** Despite noting the potential complications from LLM responses, the counselor noted the importance and demand for in-situ communication help and coaching.

**Counselor:** *Those types of strategies, you can talk a lot about it. But if you don't practice it in the moment or remind yourself of it, it's very hard to do. [...] So in the moment or close, yes, that is very, I would say, high demand and needed.*

**4.3.6 A Divergence in Participant and Practitioner Opinions.** Our findings revealed a notable divergence in the assessments of the LLM between participants and the practitioner. Participants generally found the advice to be clear, actionable, and beneficial, with it allowing for agency in navigating communication difficulties and enabling actualization of desired behavior from others. Meanwhile, the counselor expressed concerns that many responses could lead to unexpected situations or additional confusion. This disparity in response raises important questions regarding the role of LLM advice in future and implications with its use relative to *best practices*, which we address in § 6.1.

## 5 DISCUSSION

The motivation behind this work stemmed from the recent surge in popularity of large language models, and the desire to better understand the opportunities and risks from autistic workers' usage of LLMs. Our findings show unambiguously that the participants we interviewed displayed strong preferences for utilizing LLMs as social communication aids at work, signalling more widespread usage, as well as caution and warnings from a practitioner regarding its adoption. We reflect on why we may have observed the results we did, as well as address the difference in opinion between participants and practitioner, and how it portends difficulties in creating an equitable and practical LLM for providing social advice.

*Positionality Statement.* We disclose that some of the authors identify as neurodivergent, though none identify as autistic. Our perspective in approaching this work is from that of an accessibility researchers', who subscribe to the social model of disability. Our approach to this work was informed by a recognition that our participants were likely to be situated in contexts where the medical model framing and norms would be prevalent.

### 5.1 The Appeal of LLMs for an Autistic Worker and Envisioning Downstream Effects

In this section, we attempt to interpret the relationship between the myriad of factors which contribute to our participants' preferences for interacting with the LLM over the human confederate. We describe our best understanding of immediate and near-term ramifications from this realization as we forecast possible outcomes from future use.

**5.1.1 LLMs Could Be Seen As Better Than Existing Resources.** We observed that many participants reported limited or a total lack of reliable resources (§ 4.1.3), and this became one of the major bases for comparison from which participants evaluated their willingness to seek and adopt social advice from an automated agent. Given that participants were not initially made aware that only one of the two agents they interacted with was an automated agent - we noted that several participants (7 | P1 - P4, P8 - P10) expressed a desire to have *either* agent available to them. In many cases, the experiences that participants had with the LLM represented ones which held far more promise than the status quo.

As P9 remarked:

**P9:** *I think it [the LLM] is more willing to take the time to provide explanations for things I don't understand [...] which isn't the case in real-life.*

**5.1.2 Many of the Positively-Rated LLM Attributes Had Little to Do with Social Advice.** From our findings, we observe that many of the LLMs' qualities that participants reacted positively towards (§ 4.2.2) were about affective or communicative style, rather than about substantive social guidance itself. While it may not be possible to distinguish how each quality affected participants' overall attitudes, this nonetheless shows that **participants are looking for more than social advice with LLMs and are placing significant importance on the manner and context in which advice is delivered.**

Given this and our participants' perception of resource deficits (§ 4.1.3) we observe that positively rated LLM attributes have potential to give insight into addressing autistic and non-autistic dyadic or group social communication. *Specifically, behavior exhibited by LLMs gives us explicit, actionable cues on which conversations could be adapted to improve current-day communication challenges.*

Regrettably, we believe that preference for LLMs in this aspect also likely reflects a deficit in participants' current workplaces of qualities (§ 4.2.2) like open-mindedness, considerateness (in taking extra steps to make more comprehensible, legible communication), and courteousness. This conclusion is likely generalizable as it dovetails with well-founded existing research establishing greater rates of "workplace incivility" experiences by minoritized groups in the workplace [17]. We hope that positive human-LLM interactions, rather than reinforcing the commonly held ableist notion that "autistics prefer to communicate with robots," [70], can provide guidance on reframing these notions as simply *communication preferences*.

## 5.2 LLM Advice Appears Intractable, Even With Disclosure

The researchers, participants, and practitioner observed that the LLM frequently suggested advice which could be problematic for the autistic user. Some included employer-friendly language, while other strategies encouraged autistic individuals to engage in behaviors which would necessitate masking, such as maintaining eye contact, smiling, or participating in large group discussions (§ 4.3). This was perhaps to be expected, given that only a minority of participants (2) ever disclosed to the LLM that they were autistic. However, even in cases where disclosure occurred, our participants and counselor observed that these recommendations persisted, though to a lesser degree. We provide a representative example of this type of conversation in Appendix B. We see that its adjustments include the highly questionable suggestion of disclosing one's autism, which prior research has shown to be perilous [48].

Given the probabilistic and so far undefined nature of prompt engineering and variation of user input to LLM output, we would consider it unlikely that prompt engineering alone could mitigate the risks identified by the practitioner. It is clear to the authors that future LLM-based solutions to assist autistic workers' social communication ought to consider the process of more systematic and comprehensive construction for models which may be used for this purpose.

## 6 DESIGN CONSIDERATIONS FOR AUTISTIC WORKERS' USE OF LLMs

Such comprehensive efforts, which may include ensuring representativeness of autistic perspectives in training corpora, and carefully balancing the diversity of perspectives involved in value-sensitive processes such as training reward functions [65], begin to address larger sociotechnical issues which are part of the current discourse around technology and stakeholder representation. We present the following design considerations which hope provide guidance for system builders, the autism and broader accessibility community, and researchers with interest in this space.

### 6.1 LLMs for Social Advice Necessarily Involves Relative Privileging

We note that a practitioner's role is generally representative of an occupation aligned with a medicalized model of disability and autism, wherein typical goals emphasize management or mitigation of autism-related challenges. The tendency then, would be to encourage clients to adhere to neurotypical norms, sometimes at the expense of their natural inclinations. This may be, as our job coach and counselor often emphasizes, not that practitioners seek to deny autistic expression, but must nevertheless encourage a practicable level of conformity to existing workplace expectations from clients. In contrast, our participants sought an enabling support to assist with understanding an unfamiliar world which understood their perspective. Participants preferred LLM interactions despite flaws like neurotypical-centered approaches (even after disclosure), employer-centered advice, and Pollyanna-esque outlook on difficult social situations.

As such, we observe that a coach-like LLM which provides advice adhering to practitioners' best practices would not fulfill the same role as the kind of empathetic assistant in which participants would expect. In fact, as we believe that an LLM aligned with the perspective espoused by our participants will not effectively serve the broader goal of making guidance and advice from practitioners more accessible to autistic workers. Instead, it could exacerbate existing tensions between certain segments of the autistic community and practitioners. These tensions often stem from perceptions of practitioners' practices as reinforcing an oppressive status quo, forcing autistic individuals into a restrictive mold, and promoting masking.

This division begs the question of how to consider such a conflict. We believe it is imperative to discuss how to ameliorate, if not resolve the issue:

*Should the goal of LLMs as an assistive technology be to advance the interests of the disabled individual or the normative social good?*

With this question, we find that existing perspectives (§ 2.4) illustrate the gamut of different possible positions. Those identifying with the medical model of disability would argue that the goal of an LLM in this role should be to have it replicate existing options for therapy and encourage workplace assimilation. From those aligned with a standpoint theory perspective, the answer to this question may vary depending on the priority that an autistic user places between an LLM which encourages fitting in to existing workplace structures and advocating for one's own needs and expressing one's own individuality. Disability studies and critical disability advocates may argue that an LLM explicitly be constructed to prioritize and empower disabled individuals' desires first and foremost, leery of solutions which promote existing power dynamics and social structures like the existing deficit-based view of disability in many workplaces.

We believe that understanding perspectives on a values-based question such as this one is important as we increasingly utilize automated agents like LLMs for value judgments, particularly in identifying where and how such value judgments are informed in the system development process. As LLM research increasingly focuses on issues of value alignment [65], we encourage an introspection of the following question: *Whose experience should we privilege with LLMs? And to what end?*

LLM and foundation model developers are continually developing new and improved models, including those that are explicitly tuned to dispense social advice [29]. We see that deciding on which types of response to designate as the "good" or "appropriate" one is a choice which necessarily privileges the perspectives of certain parties over the other.

More concretely, we see that developing models in which practitioners are consulted for determining "goodness" would result in a relative deprioritization of the interests for the autistic user, and vice-versa. This conundrum is not easily answered. A natural and subsequent question that systems developers should ideally address whether the intention of the system is to partake in normalization of social change or to preserve the status quo of the current minoritization of autistic individuals. We note that this is likely secondary to the question of whether LLM developers will

involve stakeholders like autistic users and practitioners in LLM development, which would be a prerequisite to address this question of relative privilege.

## 6.2 Assistive, Rather Than Curative Solutions

We advocate for a focus towards developing LLMs as assistive tools that support autistic individuals — one which centers the individual’s perspective in achieving their personal and professional goals. Many prior technological interventions for autistic individuals have traditionally had a *curative* perspective, aiming to “fix” perceived deficits or challenges associated with autism. This perspective, while perhaps well-intentioned, can reinforce harmful stereotypes and stigmas, and overlook the unique strengths and perspectives that autistic individuals bring to social and workplace environments [95].

In contrast, an *assistive* approach would view LLMs as tools that can empower autistic individuals by providing support in areas where they may struggle, while also respecting and valuing their unique experiences and perspectives. This approach aligns with the Neurodiversity movement [89], which advocates for viewing autism as a natural variation of the human experience, rather than a disorder to be cured. We encourage researchers and developers to adopt this assistive perspective when designing and evaluating LLMs. This could involve focusing on how LLMs can provide practical support in areas such as social interaction or employment, while also ensuring that the tools respect and value the experiences and perspectives of autistic individuals.

## 6.3 Re-imagining Inclusive Communication Norms Involving Autistic Individuals

Our explorations with LLMs highlight concrete realizations about the potential for re-imagining inclusive communication norms involving autistic individuals. LLM’s promptness, clear formatting, and (among others), which were appreciated by participants, underscore the possibility of creating conversational spaces that are more accommodating and respectful of autistic communication styles.

This aligns with Milton’s Double Empathy perspective [56], which posits that communication difficulties between autistic and non-autistic individuals stem from mutual misunderstandings due to differing perceptions and experiences of the world. LLMs could help bridge this gap by providing neurotypicals an accessible method to adapt to a communication style that is in-line with the needs and preferences autistic individuals, thereby fostering greater mutual understanding.

We encourage further exploration on how the positive attributes of LLMs can be leveraged to promote more inclusive communication norms. Yet at the same time, we note the potential for LLMs to become a mere *translation layer* bridging neurotypicals and autistic individuals, which demands further exploration for its consequences.

## 6.4 Speculative Futuring with LLMs

If LLMs become more tightly integrated into workplace communication practices, and the phenomenon of utilizing LLMs for communications becomes commonplace, we anticipate potential difficulties

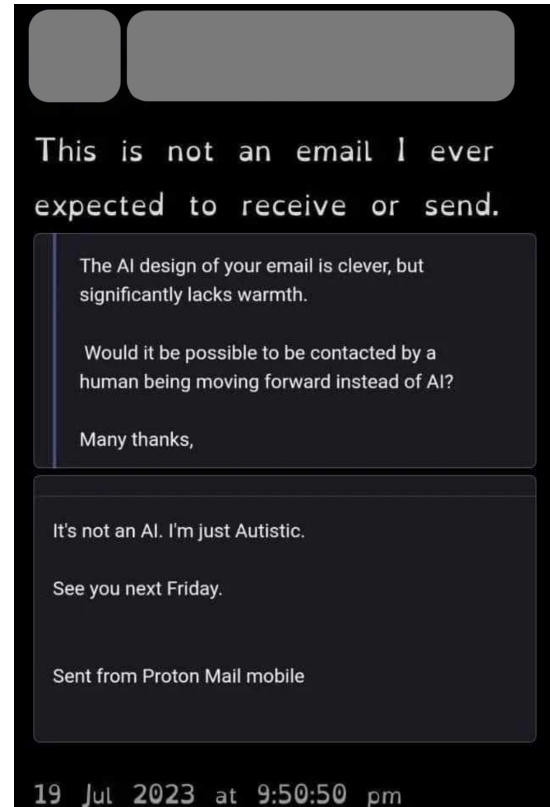


Figure 3: An example of possible social ramifications of suspected LLM use.

and confusion surrounding the ultimate role for LLMs. We are particularly concerned with the possibility that LLMs may be viewed as (or explicitly dictated by supervisors as) a necessary crutch by autistic workers in low-resource occupations.

We also note the connotation and social signalling involved with using or being suspected of utilizing an LLM. One such example, in which we provide in Figure 3, illustrates a possible consequence. In it, an autistic user remarks at the fact that they were mistaken by a different individual as having used an LLM — to which the user responds “I’m just autistic.” We believe further research is warranted to determine the permutations of social blame which those who utilize LLMs may experience. However, we do not dismiss the possibility, that ubiquitous and rampant LLM usage may dissolve those concerns altogether in a scenario where the provenance of online communicative content becomes unimportant.

## 7 LIMITATIONS

We note the following as potential factors which may affect the internal and external validity from this work, and describe relevant mitigations where warranted.

*Paprika’s Application and Operation.* We first note that the LLM used for the study, *Paprika*, may exhibit different behavior compared to an unprompted version of the GPT-4. We employed a specific prompt (§ 3.1.4) because we observed it gave responses similar

in content to the unprompted version of the GPT-4 API, while respecting guidelines on verbosity and tone to better align with the confederate’s output. We further note that as we utilized the GPT-4 API, that its cutoff date of September 2021 may have had an effect on *Paprika*’s ability to provide acceptable advice and respond to autistic perspectives.

Additionally, we implemented functionality which limited the LLMs’ verbosity and response time to be similar to that of the confederate (§ 3.1.4) with half (5 of 11) of the participant sample. When comparing ratings for participants without (first 6) and with (latter 5) adjustment (§ 4.2), we observe that participants in the latter group rated the LLM *higher*, and the confederate *lower*, than the participant group encountering *Paprika* without adjustment. Despite the different configurations of LLMs, this result appears to reinforce our perception that participants’ general preference for the LLM agent included factors outside of increased verbosity and lower response times.

*Study Design and Participant Demographics.* We note that our study was designed for 90 minutes – given this length, we acknowledge the role that participant fatigue may have played in chatbot interactions towards the end of the study, though we did not observe any significant (statistically or otherwise) differences in quantitative and qualitative results from later interactions. Relatedly, given the potential for chatbot interaction order may have (e.g. whether a participant encounters *Paprika* or *Pepper* first), we counterbalanced the order across our participant group.

We observed that our participants included 9 women and 2 men, which has a notably higher proportion of women than the diagnosed population of autistic people at large [48]. While dimensions of gender identities are known to play large roles in workplace experiences and navigating those challenges, these difficulties are amplified for underrepresented autistic adult employees, especially when finding suitable support [37, 40, 55, 58] and are often understudied in research [19]. While this aspect was not specifically explored for this work, we believe a participant group with more women could help illuminate a more diverse range of possible viewpoints than one with more men than women, considering how deeply gender identify relates to social communication issues and norms. Relatedly, we note that our participant group is almost certainly not representative of the larger autistic community owing to the high prevalence of individuals with bachelor’s or advanced degrees. Likewise, none of our participants identified as nonverbal, nor did it seem likely that a substantial part of our group had an intellectual disability.

Though our human confederate (*Pepper*) tried their best to apply a consistent approach to answering questions from participants, they did not ask follow-up questions of all participants. To mitigate this potential issue, we validated our confederate’s responses with the practitioner (§ 4.3), a professional counselor and job coach. The practitioner rated the confederate’s response as satisfactory, without caveats.

While we believe triangulation from participants’ preferences for *Paprika* (LLM) over *Pepper* (confederate) across (1) quantitative and (2) qualitative data (§ 4.2), alongside (3) participants’ desire to use *Paprika* as a sole or primary resource (§ 4.2.3), provides grounding for conclusions presented in this work, we acknowledge that

this work is exploratory in nature, and the presence of a singular confederate, along with a limited sample size ( $n = 11$ ), precludes our ability to make durable conclusions. Relatedly, we acknowledge limitations related to having a single expert review to assess the quality of our LLMs, and expect that future work into and demonstrating the applicability and hazards of LLM advice can benefit from multiple expert agreements – we see that including more job coaches could reduce the risk of overemphasis singularly on expert opinions.

As this study explores potential outcomes of autistic workers relying on publicly available general-purpose LLMs, such as ChatGPT and Gemini (née Bard), we do not make direct or specific claims about future or special-purpose models. Rather, we aim to illustrate the current and potential outcomes, opportunities, and risks from autistic workers’ use of widely-available LLMs, and to demonstrate this use as a matter of exigent concern given current the current progress in LLM development and affordability, existing and growing interest from autism-related communities, and our participants’ enthusiasm for adopting LLM advice despite its flaws.

Overall, this study examines the possible results for autistic workers who use general-purpose LLMs, like OpenAI’s ChatGPT and Google’s Gemini (née Bard). Given our study limitations and the rapidly-changing landscape of state-of-the-art language and foundation models, we do not attempt to critique or otherwise analyze current models, nor speculate about future or specialized models. Our work instead aims to highlight and centers autistic workers’ needs and desire for greater empathy and agency.

## 8 CONCLUSION

This work explores the potential and risks of autistic workers’ adoption of LLMs for social advice by simulating an experience using our own LLM and a human confederate. Our findings revealed that many of our 11 participants displayed significant affinity for LLM interactions against both the confederate and existing resources for the LLMs’ potential for increased agency in socially navigating the neurotypical world. Our participants’ enthusiastic response to our LLM was in stark contrast to our job coach’s, who urged caution and called attention to significant potential concerns from following LLM-generated advice. We show that though the LLM context may be novel, the discordance between participant and practitioner expectations towards LLMs as assistive tools echoes the impasse between medical and social models of disability, and technology and stakeholder representation – where disabled voices and experiences fall into a situation of relative privileging against normative ones. We urge the accessibility, NLP, and HCI community to adopt an assistive framing to motivate development of assistive technology which truly serve the needs and respect the autonomy of autistic individuals.

## ACKNOWLEDGMENTS

We would like to thank Amy Tavares, Frank Elavsky, Laura Dabbish, Faria Huq, Jeffrey Bigham, Brianna Blaser (and AccessComputing), the College Autism Network (CANVAS), Cella Sum, Lea Albaugh, Katie Oswald, Paulette Penzvalto, Yunzhi Li, Neeta Khanuja, and Alice Tang, for their help in ideation, development, analysis, and

recruitment for this work. We are especially grateful for our anonymous CHI reviewers for their constructive guidance and feedback through the review process.

## REFERENCES

- [1] [n. d.]. Blind - The #1 Tech Career Community – teamblind.com. <https://www.teamblind.com/>. [Accessed 05-Apr-2023].
- [2] [n. d.]. Reddit - Dive into anything – reddit.com. [https://www.reddit.com/r/autism/comments/12tk9ps/chatgpt\\_is\\_a\\_gamechanger\\_for\\_people\\_on\\_the/](https://www.reddit.com/r/autism/comments/12tk9ps/chatgpt_is_a_gamechanger_for_people_on_the/). [Accessed 13-06-2023].
- [3] Hala Annabi. 2018. The Untold Story: The Masked Experiences of Women with Autism Working in IT. (2018).
- [4] Hala Annabi and Jill Locke. 2019. A theoretical framework for investigating the context for creating employment success in information technology for individuals with autism. *Journal of Management & Organization* 25, 4 (2019), 499–515.
- [5] American Psychiatric Association, and American Psychiatric Association. 2013. *Diagnostic and statistical manual of mental disorders : DSM-5* (5th ed. ed.). American Psychiatric Association, Arlington, VA. xlv, 947 p. ; pages.
- [6] Susanna Baldwin, Debra Costley, and Anthony Warren. 2014. Employment activities and experiences of adults with high-functioning autism and Asperger's disorder. *Journal of autism and developmental disorders* 44, 10 (2014), 2440–2449.
- [7] Shaowen Bardzell and Jeffrey Bardzell. 2011. Towards a feminist HCI methodology: social science, feminism, and HCI. In *Proceedings of the SIGCHI conference on human factors in computing systems*. 675–684.
- [8] Emily M. Bender, Timnit Gebru, Angelina McMillan-Major, and Shmargaret Shmitchell. 2021. On the Dangers of Stochastic Parrots: Can Language Models Be Too Big?. In *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency* (Virtual Event, Canada) (FAccT '21). Association for Computing Machinery, New York, NY, USA, 610–623. <https://doi.org/10.1145/3442188.3445922>
- [9] Emily M Bender, Timnit Gebru, Angelina McMillan-Major, and Shmargaret Shmitchell. 2021. On the dangers of stochastic parrots: Can language models be too big?. In *Proceedings of the 2021 ACM conference on fairness, accountability, and transparency*. 610–623.
- [10] Barbara Bissonnette. 2013. *Asperger's Syndrome Workplace Survival Guide*. Jessica Kingsley, London, England.
- [11] Virginia Braun and Victoria Clarke. 2006. Using thematic analysis in psychology. *Qualitative Research in Psychology* 3, 2 (2006), 77–101.
- [12] Simon Brisenden. 1986. Independent living and the medical model of disability. *Disability, Handicap & Society* 1, 2 (1986), 173–178.
- [13] Sloane Burgess and Robert E Cimeria. 2014. Employment outcomes of transition-aged adults with autism spectrum disorders: A state of the states report. *American journal on intellectual and developmental disabilities* 119, 1 (2014), 64–83.
- [14] Moira Burke, Robert Kraut, and Diane Williams. 2010. Social use of computer-mediated communication by adults on the autism spectrum. In *Proceedings of the 2010 ACM conference on Computer supported cooperative work*. 425–434.
- [15] Courtney Byun, Piper Vasicek, and Kevin Seppi. 2023. Dispensing with Humans in Human-Computer Interaction Research. In *Extended Abstracts of the 2023 CHI Conference on Human Factors in Computing Systems*. 1–26.
- [16] Dara V Chan, Julie D Doran, and Oslly D Galobardi. 2023. Beyond friendship: The spectrum of social participation of autistic adults. *Journal of Autism and Developmental Disorders* 53, 1 (2023), 424–437.
- [17] Lilia M. Cortina, Vicki J. Magley, Jill Hunter Williams, and Regina Day Langhout. 2001. Incivility in the workplace: Incidence and impact. *Journal of Occupational Health Psychology* 6, 1 (2001), 64–80. <https://doi.org/10.1037/1076-8998.6.1.64>
- [18] Clare Cummins, Elizabeth Pellicano, and Laura Crane. 2020. Autistic adults' views of their communication skills and needs. *International journal of language & communication disorders* 55, 5 (2020), 678–689.
- [19] Joyce Davidson and Sophie Tamas. 2016. Autism and the ghost of gender. *Emotion, Space and Society* 19 (2016), 59–65.
- [20] Yao Du and Felix Juefei-Xu. 2023. Generative AI for Therapy? Opportunities and Barriers for ChatGPT in Speech-Language Therapy. (2023).
- [21] Carlton J. Fong, Joshua Taylor, Aynura Berdyeva, Amanda M. McClelland, Kathleen M. Murphy, and John D. Westbrook. 2021. Interventions for improving employment outcomes for persons with autism spectrum disorders: A systematic review update. *Campbell Systematic Reviews* 17, 3 (July 2021). <https://doi.org/10.1002/cl2.1185>
- [22] Kyle M Frost, Kathryn M Bailey, and Brooke R Ingersoll. 2019. "I just want them to see me as... me": Identity, community, and disclosure practices among college students on the autism spectrum. *Autism in Adulthood* 1, 4 (2019), 268–275.
- [23] Vinitha Gadiraju, Shaun Kane, Sunipa Dev, Alex Taylor, Ding Wang, Emily Denton, and Robin Brewer. 2023. "I wouldn't say offensive but...": Disability-Centered Perspectives on Large Language Models. In *Proceedings of the 2023 ACM Conference on Fairness, Accountability, and Transparency*. 205–216.
- [24] Peter F Gerhardt and Ilene Lainer. 2011. Addressing the needs of adolescents and adults with autism: A crisis on the horizon. *Journal of Contemporary Psychotherapy* 41 (2011), 37–45.
- [25] Parisa Ghanouni and Stephanie Quirke. 2023. Resilience and coping strategies in adults with autism spectrum disorder. *Journal of Autism and Developmental Disorders* 53, 1 (2023), 456–467.
- [26] Deepak Giri and Erin Brady. 2023. Exploring outlooks towards generative AI-based assistive technologies for people with Autism. *arXiv preprint arXiv:2305.09815* (2023).
- [27] Noah Glaser, Nigel Newbutt, Heath Palmer, and Matthew Schmidt. 2023. Video-Based Virtual Reality Technology for Autistic Users: An Emerging Technology Report. *Technology, Knowledge and Learning* 28, 2 (2023), 925–935.
- [28] Steven M Goodman, Erin Buehler, Patrick Clary, Andy Coenen, Aaron Donsbach, Tiffanie N Horne, Michal Lahav, Robert MacDonald, Rain Breaw Michaels, Ajit Narayanan, et al. 2022. LaMPost: Design and Evaluation of an AI-assisted Email Writing Prototype for Adults with Dyslexia. In *Proceedings of the 24th International ACM SIGACCESS Conference on Computers and Accessibility*. 1–18.
- [29] Nico Grant. [n. d.]. Google Tests an A.I. Assistant That Offers Life Advice – nytimes.com. <https://www.nytimes.com/2023/08/16/technology/google-ai-life-advice.html>. [Accessed 09-09-2023].
- [30] Carolyn M Grob, Dorothea C Lerman, Channing A Langlinais, and Natalie K Villante. 2019. Assessing and teaching job-related social skills to adults with autism spectrum disorder. *Journal of applied behavior analysis* 52, 1 (2019), 150–172.
- [31] Perttu Hämäläinen, Mikke Tavast, and Anton Kunnari. 2023. Evaluating Large Language Models in Generating Synthetic HCI Research Data: a Case Study. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*. 1–19.
- [32] Aimi Hamraie and Kelly Fritsch. 2019. Crip technoscience manifesto. *Catalyst: Feminism, Theory, Technoscience* 5, 1 (2019), 1–33.
- [33] Erin Harmuth, Emerald Silletta, Anthony Bailey, Tracy Adams, Charlotte Beck, and Skye Pamela Barbic. 2018. Barriers and facilitators to employment for adults with autism: A scoping review. *Annals of International Occupational Therapy* 1, 1 (2018), 31–40.
- [34] Drew Harwell, Nitasha Tiku, and Will Oremus. 2022. Stumbling with their words, some people let AI do the talking. <https://www.washingtonpost.com/technology/2022/12/10/chatgpt-ai-helps-written-communication/>
- [35] Elizabeth McGhee Hassrick, Laura Graham Holmes, Collette Sosnowy, Jessica Walton, and Kathleen Carley. 2021. Benefits and risks: A systematic review of information and communication technology use by autistic people. *Autism in Adulthood* 3, 1 (2021), 72–84.
- [36] Gillian R Hayes, V Erick Custodio, Oliver L Haimson, Kathy Nguyen, Kathryn E Ringland, Rachel Rose Ulgado, Aaron Waterhouse, and Rachel Weiner. 2015. Mobile video modeling for employment interviews for individuals with autism. *Journal of Vocational Rehabilitation* 43, 3 (2015), 275–287.
- [37] Susan M Hayward, Keith R McVilly, and Mark A Stokes. 2018. "Always a glass ceiling." Gender or autism; the barrier to occupational inclusion. *Research in Autism Spectrum Disorders* 56 (2018), 50–60.
- [38] Emily Heng, Marc J Lanovaz, and Alexia Beauregard. 2021. Research on technological interventions for young children with autism spectrum disorders: A scoping review. *Review journal of autism and developmental disorders* 8 (2021), 253–263.
- [39] Philippa L Howard and Felicity Sedgewick. 2021. 'Anything but the phone!': Communication mode preferences in the autism community. *Autism* 25, 8 (2021), 2265–2278.
- [40] Yunhe Huang, Samuel RC Arnold, Kitty-Rose Foley, and Julian N Trollor. 2022. Experiences of Support Following Autism Diagnosis in Adulthood. *Journal of Autism and Developmental Disorders* (2022), 1–14.
- [41] Karen R Johnson, Demetria Ennis-Cole, and Matthew Bonhamgregory. 2020. Workplace success strategies for employees with autism spectrum disorder: A new frontier for human resource development. *Human Resource Development Review* 19, 2 (2020), 122–151.
- [42] Enkelejda Kasneci, Kathrin Seßler, Stefan Küchemann, Maria Bannert, Daryna Dementieva, Frank Fischer, Urs Gasser, Georg Groh, Stephan Günemann, Eyke Hüllermeier, et al. 2023. ChatGPT for good? On opportunities and challenges of large language models for education. *Learning and individual differences* 103 (2023), 102274.
- [43] Lorcan Kenny, Caroline Hattersley, Bonnie Molins, Carole Buckley, Carol Povey, and Elizabeth Pellicano. 2015. Which terms should be used to describe autism? Perspectives from the UK autism community. *Autism* 20, 4 (July 2015), 442–462. <https://doi.org/10.1177/1362361315588200>
- [44] Neha U Keshav, Joseph P Salisbury, Arshya Vahabzadeh, and Ned T Sahin. 2017. Social communication coaching smartglasses: Well tolerated in a diverse sample of children and adults with autism. *JMIR mHealth and uHealth* 5, 9 (2017), e8534.
- [45] Anna Krzeminska, Robert D. Austin, Susanne M. Bruyère, and Darren Hedley. 2019. The advantages and challenges of neurodiversity employment in organizations. *Journal of Management and Organization* 25, 4 (2019), 453–463. <https://doi.org/10.1017/jmo.2019.58>



- [46] Hanna Lagerqvist. 2020. “Seen it, enjoyed it, bye!”: Using and interacting on the Instagram platform from the perspective of autistic users.
- [47] Nathan Lambert, Luis Castricato, Leandro von Werra, and Alex Havrilla. 2022. Illustrating reinforcement learning from human feedback (RLHF). <https://huggingface.co/blog/rlhf>
- [48] Debra J. Leven. 2020. *Invisible Social Identity in the Workplace*. Ph.D. Dissertation. Northeastern University.
- [49] Sally Lindsay, Victoria Osten, Mana Rezai, and Sunny Bui. 2021. Disclosure and workplace accommodations for people with autism: A systematic review. *Disability and rehabilitation* 43, 5 (2021), 597–610.
- [50] Beatriz López. 2022. Commentary on Autism and the double-empathy problem: Implications for development and mental health. *The British Journal of Developmental Psychology* 40, 2 (2022), 368.
- [51] Kelly Mack, Emma McDonnell, Dhruv Jain, Lucy Lu Wang, Jon E. Froehlich, and Leah Findlater. 2021. What Do We Mean by “Accessibility Research”? A Literature Survey of Accessibility Papers in CHI and ASSETS from 1994 to 2019. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (Yokohama, Japan) (CHI '21). Association for Computing Machinery, New York, NY, USA, Article 371, 18 pages. <https://doi.org/10.1145/3411764.3445412>
- [52] Jennifer Mankoff, Gillian R. Hayes, and Devva Kasnitz. 2010. Disability Studies as a Source of Critical Inquiry for the Field of Assistive Technology. In *Proceedings of the 12th International ACM SIGACCESS Conference on Computers and Accessibility* (Orlando, Florida, USA) (ASSETS '10). Association for Computing Machinery, New York, NY, USA, 3–10. <https://doi.org/10.1145/1878803.1878807>
- [53] Deborah Marks. 1997. Models of disability. *Disability and rehabilitation* 19, 3 (1997), 85–91.
- [54] Matthew B Miles, A Michael Huberman, and Johnny Saldana. 2019. *Qualitative Data Analysis* (4 ed.). SAGE Publications, Thousand Oaks, CA.
- [55] Kaitlin H Koffer Miller, Dylan S Cooper, Wei Song, and Lindsay L Shea. 2022. Self-reported service needs and barriers reported by autistic adults: differences by gender identity. *Research in Autism Spectrum Disorders* 92 (2022), 101916.
- [56] Damian EM Milton. 2012. On the ontological status of autism: The ‘double empathy problem’. *Disability & society* 27, 6 (2012), 883–887.
- [57] Michael Muller, Lydia B Chilton, Anna Kantosalo, Charles Patrick Martin, and Greg Walsh. 2022. GenAI/CHI: Generative AI and HCI. In *CHI Conference on Human Factors in Computing Systems Extended Abstracts*. 1–7.
- [58] Wasan Nagib and Robert Wilton. 2020. Gender matters in career exploration and job-seeking among adults with autism spectrum disorder: evidence from an online community. *Disability and rehabilitation* 42, 18 (2020), 2530–2541.
- [59] David B. Nicholas, Darren Hedley, Jena K. Randolph, Dora M. Raymaker, Scott M. Robertson, and Jonathan Vincent. 2019. An Expert Discussion on Employment in Autism. *Autism in Adulthood* 1, 3 (Sept. 2019), 162–169. <https://doi.org/10.1089/aut.2019.29003.djn>
- [60] David B Nicholas, Wendy Mitchell, Carolyn Dudley, Margaret Clarke, and Rosslynn Zulla. 2018. An ecosystem approach to employment and autism spectrum disorder. *Journal of autism and developmental disorders* 48, 1 (2018), 264–275.
- [61] Christine L Nittrouer, Evan E Dean, and Karrie A Shogren. 2022. Career progression: Strategies used by self-advocates with ASD and their allies. In *Generation A: Perspectives on Special Populations and International Research on Autism in the Workplace*. Emerald Publishing Limited, 45–63.
- [62] Mike Oliver. 2013. The social model of disability: Thirty years on. *Disability & society* 28, 7 (2013), 1024–1026.
- [63] OpenAI. 2022. Introducing chatgpt. <https://openai.com/blog/chatgpt>
- [64] OpenAI. 2023. GPT-4 Technical Report. <https://doi.org/10.48550/ARXIV.2303.08774>
- [65] Long Ouyang, Jeffrey Wu, Xu Jiang, Diogo Almeida, Carroll Wainwright, Pamela Mishkin, Chong Zhang, Sandhini Agarwal, Katarina Slama, Alex Ray, John Schulman, Jacob Hilton, Fraser Kelton, Luke Miller, Maddie Simens, Amanda Askell, Peter Welinder, Paul F Christiano, Jan Leike, and Ryan Lowe. 2022. Training language models to follow instructions with human feedback. In *Advances in Neural Information Processing Systems*, S. Koyejo, S. Mohamed, A. Agarwal, D. Belgrave, K. Cho, and A. Oh (Eds.), Vol. 35. Curran Associates, Inc., 27730–27744. [https://proceedings.neurips.cc/paper\\_files/paper/2022/file/b1efde53be364a73914f58805a001731-Paper-Conference.pdf](https://proceedings.neurips.cc/paper_files/paper/2022/file/b1efde53be364a73914f58805a001731-Paper-Conference.pdf)
- [66] Sundar Pichai. 2023. An important next step on our ai journey. <https://blog.google/technology/ai/bard-google-ai-search-updates/>
- [67] Christina H. Rebholz. 2012. *Life in the Uncanny Valley: Workplace Issues for Knowledge Workers on the Autism Spectrum*. Ph.D. Dissertation. Antioch University. <http://aura.antioch.edu/etds/29>
- [68] Meredith Ringel Morris, Carrie J Cai, Jess Holbrook, Chinmay Kulkarni, and Michael Terry. 2023. The Design Space of Generative Models. *arXiv e-prints* (2023), arXiv–2304.
- [69] Milly Romeijn-Stout, Mina Tari, and Hala Annabi. [n. d.]. Examining Bias in Screening Interviews Towards Individuals with Autism in the Tech Industry. ([n. d.]).
- [70] Helen Rottier, Ben Pflingston, and Joshua Guberman. 2022. Ghosts, mice, and robots: DisAppearing the autistic person. *DisAppearing: Encounters in disability studies*. Toronto: Canadian Scholars (2022), 93–105.
- [71] Anne M. Roux, Paul T. Shattuck, Jessica E. Rast, Julianna A. Rava, and Kristy A. Anderson. 2015. National Autism Indicators Report: Transition into Young Adulthood. *Life Course Outcomes Research Program* (2015). <https://doi.org/10.1007/s10803-010-1070-3>
- [72] Maureen M Schepis, Dennis H Reid, Michael M Behrmann, and Kelly A Sutton. 1998. Increasing communicative interactions of young children with autism using a voice output communication aid and naturalistic teaching. *Journal of applied behavior analysis* 31, 4 (1998), 561–578.
- [73] Ralf W Schlosser and Rajinder K Koul. 2015. Speech output technologies in interventions for individuals with autism spectrum disorders: A scoping review. *Augmentative and Alternative Communication* 31, 4 (2015), 285–309.
- [74] Melissa Scott, Marita Falkmer, Sonya Girdler, and Torbjörn Falkmer. 2015. Viewpoints on Factors for Successful Employment for Adults with Autism Spectrum Disorder. *PLOS ONE* 10, 10 (Oct. 2015), e0139281. <https://doi.org/10.1371/journal.pone.0139281>
- [75] Tom Shakespeare et al. 2006. The social model of disability. *The disability studies review* 2 (2006), 197–204.
- [76] Samantha Spengler. 2023. For Some Autistic People, ChatGPT Is a Lifeline – wired.com. <https://www.wired.com/story/for-some-autistic-people-chatgpt-is-a-lifeline/>. Accessed: 04 Jun 2023].
- [77] Lucy Suchman. 2016. Organizing alignment: the case of bridge-building. In *Knowing in organizations: a practice-based approach*. Routledge, 187–203.
- [78] Kathryn A Szechy, Pamela D Turk, and Lisa A O'Donnell. 2023. Autism and Employment Challenges: The Double Empathy Problem and Perceptions of an Autistic Employee in the Workplace. *Autism in Adulthood* (2023).
- [79] Hannah Taylor, Christopher J Hand, Hannah Howman, and Ruth Filik. 2022. Autism, attachment, and alexithymia: Investigating emoji comprehension. *International Journal of Human-Computer Interaction* (2022), 1–14.
- [80] Julie Louds Taylor and Marsha Mailick Seltzer. 2010. Employment and Post-Secondary Educational Activities for Young Adults with Autism Spectrum Disorders During the Transition to Adulthood. *Journal of Autism and Developmental Disorders* 41, 5 (jul 2010), 566–574. <https://doi.org/10.1007/s10803-010-1070-3>
- [81] Sandra Thompson-Hodgetts, Chantal Labonte, Rinita Mazumder, and Sharon Phelan. 2020. Helpful or harmful? A scoping review of perceptions and outcomes of autism diagnostic disclosure to others. *Research in Autism Spectrum Disorders* 77 (2020), 101598.
- [82] Katie Togher and Sarah Jay. 2023. Disclosing an autism diagnosis: A social identity approach. *Autism Research* (2023).
- [83] u/ComprehensiveTime634. 2023. GPT for Interpreting Vague Instructions at Work. [https://reddit.com/r/autism/comments/182zwme/gpt\\_for\\_interpreting\\_vague\\_instructions\\_at\\_work/](https://reddit.com/r/autism/comments/182zwme/gpt_for_interpreting_vague_instructions_at_work/). Accessed: 24 Nov 2023.
- [84] u/hadleyajohnston. 2023. CHATGPT as a Tool for Autistics!! [https://www.reddit.com/r/autism/comments/10wv1rl/chatgpt\\_as\\_a\\_tool\\_for\\_autistics/](https://www.reddit.com/r/autism/comments/10wv1rl/chatgpt_as_a_tool_for_autistics/). Accessed: 8 Feb 2023.
- [85] u/lilgreycat. 2023. What Do You Use Chatgpt For? [https://www.reddit.com/r/autism/comments/13tlhop/what\\_do\\_you\\_use\\_chatgpt\\_for/](https://www.reddit.com/r/autism/comments/13tlhop/what_do_you_use_chatgpt_for/). Accessed: 27 May 2023.
- [86] u/SACBH. 2023. Chatgpt Is a Gamechanger for People on the Spectrum. [https://reddit.com/r/autism/comments/12tk9ps/chatgpt\\_is\\_a\\_gamechanger\\_for\\_people\\_on\\_the/](https://reddit.com/r/autism/comments/12tk9ps/chatgpt_is_a_gamechanger_for_people_on_the/). Accessed: 21 Apr 2023.
- [87] Stephanie Valencia, Richard Cave, Krystal Kallarackal, Katie Seaver, Michael Terry, and Shaun K Kane. 2023. “The less I type, the better”: How AI Language Models can Enhance or Impede Communication for AAC Users. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*. 1–14.
- [88] Thijs Waardenburg, Niels van Huizen, Jelle van Dijk, Koen Dortmans, Maurice Magnée, Wouter Staal, Jan-Pieter Teunisse, and Mascha van der Voort. 2022. Design your life: user-initiated design of technology to empower autistic young adults. *Journal of Enabling Technologies* 16, 3 (2022), 172–188.
- [89] Nick Walker. 2021. *Neuroqueer heresies*. Autonomous Press, Bondi Junction, NSW, Australia.
- [90] Emmanuelle Walkowiak. 2021. Neurodiversity of the workforce and digital transformation: The case of inclusion of autistic workers at the workplace. *Technological Forecasting and Social Change* 168 (2021), 120739. <https://doi.org/10.1016/j.techfore.2021.120739>
- [91] Lisa Walsh, Sinead Lydon, and Olive Healy. 2014. Employment and vocational skills among individuals with autism spectrum disorder: Predictors, impact, and interventions. *Review Journal of Autism and Developmental Disorders* 1 (2014), 266–275.
- [92] Manhua Wang and Myoungsoon Jeon. 2023. Assistive Technology for Adults on the Autism Spectrum: A Systematic Survey. *International Journal of Human-Computer Interaction* (2023), 1–20.
- [93] Justin D Weisz, Michael Muller, Jessica He, and Stephanie Houde. 2023. Toward general design principles for generative AI applications. *arXiv preprint arXiv:2301.05578* (2023).
- [94] Meredith Whittaker, Kate Crawford, Roel Dobbe, Genevieve Fried, Elizabeth Kazianas, Varoon Mathur, Sarah Myers West, Rashida Richardson, Jason Schultz, Oscar Schwartz, et al. 2018. *AI now report 2018*. AI Now Institute at New York University New York.

- [95] Rua Mae Williams, Louanne Boyd, and Juan E Gilbert. 2023. Counterinterventions: a reparative reflection on interventionist HCI. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*. 1–11.
- [96] Andreas Witzel. 2000. The Problem-centered Interview. *Forum Qualitative Sozialforschung / Forum: Qualitative Social Research* Vol 1 (2000), Methodical and Empirical Examples. <https://doi.org/10.17169/FQS-1.1.1132>
- [97] Lian Zhang. 2018. *Design and Applications of Intelligent Human-Computer Interaction Systems for Autism Spectrum Disorder intervention*. Ph. D. Dissertation.
- [98] Annuska Zolyomi, Andrew Begel, Jennifer Frances Waldern, John Tang, Michael Barnett, Edward Cutrell, Daniel McDuff, Sean Andrist, and Meredith Ringel Morris. 2019. Managing stress: the needs of autistic adults in video calling. *Proceedings of the ACM on Human-Computer Interaction* 3, CSCW (2019), 1–29.

## A CODES FROM THEMATIC ANALYSIS

<b>Communication Challenges</b>	<b>Support and Resources</b>	<b>Job Environment</b>	<b>Attitudes and Expectations towards LLMs (pre-study)</b>	<b>Attitudes and Expectations towards LLMs (post-study)</b>
Communication norms (implicit / explicit)	Disability	Job Types	Hopes for Chatbots	Benefits to Chatbots
Supervisor or coworker feedback	Accommodations	Communication modes	Problems with Existing Chatbots	Chatbot is better than...
Response to supervisor or coworker feedback	Disclosure	Sensory overstimulation	Experience with chatbots	Envisioning the Role of an LLM
Metacognitive response to feedback	Resources (Workplace / Vocational / Personal)	Situational hazards		Concerns about LLM interaction
Managing interpersonal relationships	Helpfulness of Resources			Hopes for the future (of their work/themselves)
Downstream impact of challenges on Work / Life	Considerations for asking for help			Desires for Chatbot Answers
				Positive response towards LLM
				Positive response towards human
				Negative response towards LLM
				Negative response towards human

**Table 4: Themes generated from qualitative thematic analysis.**

## B REPRESENTATIVE EXAMPLE OF AN LLM ADJUSTING TO AUTISM DISCLOSURE

This example was a simulation run with the LLM-based chatbot *Paprika* involving a question asked previously by a participant, now with a follow-up disclosure, where we observe some LLM corrections to recommended masking behaviors.

**Q: I'm about to go to a conference. How do I go about forming connections?**

**A:** *Making connections at a conference can be a rewarding experience. Here are some tips to help you make friends and build connections:*

- (1) *Be approachable: Smile and maintain open body language. This will make you appear friendly and approachable to others.*
- (2) *Start conversations: Don't be afraid to initiate conversations with people around you. You can start by asking about their work, their interests, or their reason for attending the conference.*
- (3) *Use icebreakers: If you're unsure how to start a conversation, use icebreakers such as asking about their favorite presentation, what they're looking forward to, or if they've attended the conference before.*
- (4) *Attend networking events: Many conferences have social events and networking opportunities. Attend these events and actively participate in conversations.*
- (5) *Join group discussions: During breaks or meals, join group discussions and contribute your thoughts and opinions.*
- (6) *Offer help or assistance: If you see someone struggling with something or in need of help, offer your assistance. This can be a great way to start a conversation and make a positive impression.*
- (7) *Be a good listener: When engaging in conversations, make sure to listen actively and show genuine interest in what the other person is saying. This will make them feel valued and more likely to befriend you.*
- (8) *Exchange contact information: When you meet someone you'd like to stay in touch with, exchange business cards or connect on social media platforms like LinkedIn.*
- (9) *Follow up after the conference: After the conference, send a message or email to the people you connected with, expressing how much you enjoyed meeting them and discussing shared interests or potential collaborations.*
- (10) *Be yourself: Lastly, be genuine and authentic in your interactions. People appreciate sincerity and are more likely to connect with you if they feel you're being honest and true to yourself.*

**Q: I'm actually autistic, so much of what you're suggesting is difficult for me. How would you change your advice?**

**A:** *I understand that social situations can be more challenging for individuals with autism. Here are some adjusted tips to help you form connections at a conference:*

- (1) *Prepare in advance: Before attending the conference, research the event, the speakers, and the topics being discussed. This can help you feel more confident and comfortable when engaging in conversations.*
  - (2) *Focus on your interests: Identify the sessions and topics that interest you the most. People attending these sessions are likely to share your interests, making it easier to connect with them.*
  - (3) *Use visual cues: Wear something that represents your interests, such as a pin or a shirt with a logo related to your field. This can serve as a conversation starter and help others approach you.*
  - (4) *Find a buddy: If possible, attend the conference with a friend or colleague who can help facilitate introductions and conversations.*
  - (5) *Use technology: Use social media platforms like Twitter or a conference-specific app to engage with other attendees and share your thoughts on presentations.*
  - (6) *Choose smaller sessions: Opt for smaller, more intimate sessions or workshops where it might be easier for you to engage with others and ask questions.*
  - (7) *Prepare conversation starters: Before the event, think of a few conversation starters or questions related to the conference topics. This can help reduce anxiety when initiating conversations.*
  - (8) *Take breaks: Conferences can be overwhelming, especially for individuals with autism. Schedule regular breaks to give yourself time to recharge and process the information.*
  - (9) *Set realistic goals: Instead of aiming to make friends with everyone at the conference, set a goal to connect with a few people who share your interests.*
  - (10) *Be open about your autism: If you feel comfortable, you can mention your autism to others when engaging in conversations. This can help them understand your communication style and make them more receptive to your needs.*
- Remember, it's okay to feel overwhelmed or anxious in social situations. Focus on what you can do and give yourself credit for stepping out of your comfort zone.*