Expecting Too Much, Getting Too Little: Exploring the Challenges and Design Opportunities of Asynchronous AI Interviewers

MD NAZMUS SAKIB, University of Maryland, Baltimore County, USA
NAGA MANOGNA RAYASAM, University of Maryland, Baltimore County, USA
SANORITA DEY, University of Maryland, Baltimore County, USA

Organizations use asynchronous AI interview systems to efficiently manage large applicant pools, enabling quick and uniform evaluations. However, concerns remain about their impact on user agency and the lack of personalization applicants experience with these systems. Although efforts have been made to humanize the interview process, users' expectations are often unmet, especially when compared to the promises made by these systems. To examine how applicants perceive and experience these tools, particularly as they become more familiar with widely accessible AI technologies, we conducted a two-phase study. The first phase involved an analysis of 11 subreddit discussions on interview experiences with asynchronous AI interviewers, followed by a semi-structured interview study with 17 participants. Qualitative analysis revealed key issues such as mismatched expectations, amplified by organizational rhetoric and applicant expectations shaped by experiences with Large Language Models (LLMs). These factors shaped participants' sense of agency and trust, often leading to workarounds and deceptive practices. In the follow-up study, we designed an interface with two features, response variants and feedback variants, and evaluated it across six groups (N = 180, 30 participants each) to assess whether these features support users' sense of agency, competence, and relatedness. Our analysis suggests that even subtle design changes can enhance user autonomy and that carefully designed feedback can provide meaningful support in high-stakes interview contexts.

CCS Concepts: • **Human-centered computing** → **Empirical studies in collaborative and social computing**; **User studies**.

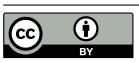
Additional Key Words and Phrases: Asynchronous AI Interviewer, AI-mediated Recruitment, Reddit, Large Language Models, LLM, Prompt Engineering, Self-Determination Theory

ACM Reference Format:
Md Nazmus Sakib, Naga Manogna Rayasam, and Sanorita Dey. 2026. Expecting Too Much, Getting Too Little: Exploring the Challenges and Design Opportunities of Asynchronous AI Interviewers. *Proc. ACM Hum.-Comput. Interact.* 10, 6, Article CSCW083 (October 2026), 40 pages. https://doi.org/10.1145/3816931

1 Introduction

The adoption of Artificial Intelligence (AI) in recruitment represents a major transformation in human resource management (HRM), aimed at increasing efficiency and fairness in candidate evaluation. AI systems now perform functions traditionally handled by recruiters, including parsing resumes, analyzing candidate responses, and interpreting multimodal data such as text, facial cues, and voice [35]. Applications span semantic matching [156], interview analysis [116], performance and attrition prediction [141], fair ranking [50], gamified assessments [71], and biometric screening

Authors' Contact Information: Md Nazmus Sakib, msakib1@umbc.edu, University of Maryland, Baltimore County, Baltimore, Maryland, USA; Naga Manogna Rayasam, tp34657@umbc.edu, University of Maryland, Baltimore County, Baltimore, Maryland, USA; Sanorita Dey, sanorita@umbc.edu, University of Maryland, Baltimore County, Baltimore, Maryland, USA.



ACM 2573-0142/2026/10-ARTCSCW083
https://doi.org/10.1145/3816931

[118]. While AI-driven technologies offer speed and scalability, they raise concerns around bias, limited contextual understanding, and a lack of transparency [16, 52]. In response to rising user expectations, many platforms now integrate emotional AI and Large Language Models (LLMs) to create more interactive interview experiences, both synchronous and asynchronous [34, 120]. However, the shift toward AI-driven evaluations has raised critical questions about fairness, candidate autonomy, and the erosion of human judgment in high stakes decision-making.

Building on these concerns, the CSCW community has critically explored the social and ethical dimensions of AI in recruitment. Emotion AI has been a key focus, with studies linking applicant perceptions to various forms of injustice [115], revealing inflated claims about solving hiring "inaccuracy" and "inauthenticity" [120], and exposing ethical risks in workplace patents that disproportionately affect low-power workers [19]. Emotional surveillance has also been shown to erode trust and wellbeing [7, 31]. Social class bias persists, as hiring practices often favor communication styles associated with upper-middle-class backgrounds [27, 28], while referrals remain more effective than algorithmic assessments [10]. Despite interest in fairness, practitioners struggle to apply formal fairness metrics in practice [126].

Recent research has shown that AI-mediated interviews shape not only candidate behavior but also perceptions of fairness, control, and authenticity. In asynchronous formats, automated evaluations tend to reduce deceptive behavior, yet they also restrict candidates' ability to express themselves fully, making the interaction feel constrained and impersonal [134]. Many applicants perceive algorithmic hiring as less fair than human or hybrid processes, even when they receive favorable outcomes, due to a lack of recognition of their individuality and context [89]. These concerns can arise even before the interview begins; job advertisements that mention AI-led interviews often reduce organizational appeal and applicants' intention to apply [147]. This discomfort becomes more pronounced in fully automated, high stakes settings, where candidates report lower feelings of control, fairness, and social presence compared to human-led formats [85]. Although AI is often associated with efficiency and objectivity, many candidates still prefer human involvement for its capacity to show empathy and adapt to context [101]. Design elements such as warm avatars and informative feedback have been shown to improve perceptions of fairness and interpersonal respect [100]. On a behavioral level, AI evaluation has been found to improve speech rate and uncertainty, suggesting that candidates adjust their communication style in response to reduced social cues and feedback [94]. While asynchronous interviews generally lead to less deceptive impression management than synchronous ones, both AI and human raters struggle to detect such behaviors accurately [134]. The use of AI-driven systems as interviewers in virtual environments also introduces concerns around trust, authenticity, and emotional connection [58]. Moreover, scholars have questioned whether AI assessments can adequately account for socially constructed traits, raising concerns about how identity and intent are constrained in such systems [4].

While prior work has shed light on candidate perceptions and design features of AI-driven interviews, important gaps remain. Most studies rely on controlled settings, offering limited insight into how job seekers engage with these systems in real-world contexts. Moreover, few examined how expectations have shifted with the emergence of LLMs. With the aim to address these gaps, we ask the following questions:

- **RQ_1:** How do candidates experience and perceive asynchronous AI-driven interviews in the current landscape of widely accessible AI technologies? What factors contribute to the gap between the intended benefits of these AI-driven systems and the experiences reported by interviewees?

- **RQ_2:** How might design considerations, informed by challenges identified in candidate experiences with AI-driven interviews, improve perceptions of control, support, and overall interaction quality?

To answer these questions, we conducted two studies. For RQ_1, we analyzed Reddit discourse to capture unfiltered perspectives, followed by in-depth interviews to understand how candidates navigate, adapt to, and make sense of AI driven interviews in the context of widely accessible AI technologies. We analyzed ~18K Reddit data collected from 11 subreddits, related to applicants' experience of interacting with AI-driven interviewers. Afterwards, we conducted a semi-structured interview study with 17 participants. The interview study allowed us to go deeper into the challenges and concerns that we found from the Reddit analysis. Our qualitative analysis revealed that applicants often entered AI driven interviews with high expectations shaped by language or terminologies used by organizations and public familiarity with LLMs, only to be disappointed by conversational yet impersonal and rigid experiences. Some used workarounds to bypass the system, citing perceived unfairness. Neurodivergent individuals reported exclusion and shared coping strategies. While consistency and efficiency were acknowledged, acceptance was largely driven by reluctance rather than trust. The lack of transparency from hiring companies, absence of acknowledgment, and overall undesired experiences left participants feeling devalued, undermining both their sense of autonomy and self-efficacy.

Based on insights from the first study, which showed that applicants often felt devalued by the lack of personalized acknowledgment and perceived limited control over how they presented themselves, we addressed RQ_2 by exploring two design considerations. They focused on providing personalized feedback and increasing applicants' flexibility in responding. We subsequently designed an interview interface incorporating two feature types, each offered in three variants: response variants (RVs), which included only re-record, only edit, or both options; and feedback variants (FVs), which included motivational, evaluative, or combined feedback. We evaluated these variants using the lens of Self-Determination Theory to examine how they might influence users' sense of agency, competence, and relatedness, and in turn their performative confidence. We conducted a user study across six groups (N = 180, 30 participants each) and analyzed their post-task survey responses using a mixed-method approach. Our results suggest that response editing improved autonomy and competence, while re-recording was useful but sometimes frustrating, and combining both options introduced decision friction. Motivational feedback felt supportive, evaluative feedback was helpful but distant, and their combination sometimes lacked clarity, highlighting the need for thoughtful and flexible design.

The main contributions of our work are as follows:

- Presenting large-scale analysis of 11 subreddit discourse with AI-driven asynchronous interviews to uncover how job seekers experience and adapt to AI-driven interviews in real-world contexts.
- Identifying key experiential gaps arising from organizational language, familiarity with contemporary AI technologies, and system limitations, with particular attention to how these affect autonomy, trust, and inclusion.
- Designing an AI-driven interview interface and evaluating it through the lens of Self-Determination Theory to examine how response and feedback features can improve user experience by supporting agency, competence, and emotional engagement.
- Extending the CSCW research direction on the adoption of new technologies by focusing on how people's perceptions of social and organizational factors influence both technology adoption and the development of critical workarounds.

We organize the paper as follows. In Section 2, we review related work. Section 3 and Section 4 describe the methodology and findings of Study 1, while Section 5 and Section 6 present the method and results of Study 2. We conclude by discussing the broader implications of our work for the CSCW community, AI mediated hiring practices, and the design of human-centered AI (HCAI) systems.

2 Related Work

2.1 Reframing Interviews in the Age of AI

2.1.1 From Dyadic Interaction to Algorithmic Assessment. The interview process has long been a subject of HCI research, particularly concerning how relational dynamics shape candidate experiences [142]. Human interviews involve co-regulated verbal and nonverbal cues, with applicants expected to demonstrate both technical and interpersonal skills [47]. However, applicants often misjudge their performance, struggle to manage anxiety-induced behaviors, and fail to mask unconscious nonverbal cues [39, 129]. AI-driven interviews, especially in asynchronous settings, have fundamentally altered this interaction. Applicants must now perform without real-time feedback or mutual adaptation, leading to heightened uncertainty, faster speech, and more constrained responses [4, 89].

Many AI interview platforms now incorporate emotion AI to assess facial expressions, voice tone, and behavioral micro-features [103, 121]. These assessments are often based on invalidated or pseudo-scientific assumptions [120], and their use has raised critical concerns about reinforcing biases and penalizing marginalized candidates for behaviors outside dominant affective norms [102]. Studies show that such systems can undermine applicant autonomy, especially when candidates cannot contest or understand the evaluation process [7, 31, 115], leaving applicants feeling dehumanized by opaque systems that offer little room for context or individuality. Particularly for neurodivergent or culturally diverse users, emotion-based evaluations risk flattening complex identities and failing to accommodate variation in expressive behavior [27, 94].

2.1.2 Perceptions of Fairness, Trust, and Transparency. Despite claims of improved efficiency and fairness, AI hiring systems often evoke skepticism and distrust. Applicants express concern over algorithmic opacity [11, 28], limited legal accountability [23], and the absence of explainability or recourse mechanisms [19, 42]. While some users perceive AI as less biased than humans [79], many remain uncomfortable with high-stakes decisions made without human oversight. Hybrid formats that combine AI assessments with human involvement are generally preferred [4] especially for roles requiring empathy or contextual understanding [101]. Notably, even including the word "AI" in job postings can lower candidate interest, suggesting deep-seated apprehension about algorithmic coldness and lack of fairness [109, 147]. Recent advances such as LLM-based interviewers aim to simulate more natural interaction [48, 93], but they raise similar concerns regarding trust, output reliability, and how well they represent human values [94, 110]. Applicants still feel the lack of transparency on how their inputs are processed, interpreted, and scored.

2.1.3 Behavioral Responses and Design Directions. AI interviews influence candidate behavior in distinct ways: participants tend to speak less, show fewer expressions, and reduce impression management, largely due to stress, uncertainty, and lack of reciprocal cues [84, 94, 147]. While interface enhancements such as warm avatars, motivational prompts, or expressive feedback may improve perceptions of fairness and reduce anxiety [51, 79, 100], these adjustments often remain cosmetic. Without transparency, interpretability, or space for meaningful self-representation, candidates continue to face emotional labor, disempowerment, and limited control over how they are evaluated.

As platforms like HireVue[1], Quinncia[2], and other asynchronous tools became widely adopted, researchers advocate for candidate-centered approaches, emphasizing transparency, contestability, procedural fairness, and support for individual voice [64, 92, 135]. However, many AI systems still neglect context-sensitive traits such as resilience or lived experience, reinforcing behavioral norms that privilege dominant cultural expressions while marginalizing applicants from diverse socio-economic and cultural backgrounds [27, 102, 121].

2.2 Supporting Agency and Engagement through Design

Self-Determination Theory (SDT) has become a foundational lens in HCI for understanding how systems can support users' basic psychological needs: autonomy, competence, and relatedness. In the context of conversational agents, SDT-informed design emphasizes personalization, flexible dialogue, and control over data to support autonomy and competence, though effectively providing relatedness remains challenging [154]. Broader critiques argue that SDT is often applied superficially; particularly in gaming and educational technologies where it serves more as a design heuristic than a robust theoretical framework [12, 37, 143]. Empirical research has shown that SDT-informed features, such as setting clear goals and incorporating perspective, enhance engagement and learning in instructional systems, and that positive user experiences can arise even from extrinsically motivated use, depending on how systems frame interactions [14]. However, attempts to establish connection through design may unintentionally compromise autonomy when relational dynamics are not carefully managed [146].

In proactive voice assistants, users respond more positively to systems that clearly communicate intent and respect user choice, underscoring the need for transparent, autonomy-supportive interactions even in anticipatory technologies [65]. AI-mediated feedback environments also demonstrate that the framing and tone of feedback significantly affect user engagement. Narrative empathy, co-constructed responses, and the presence of a perceived partnership all improve feedback acceptance and sustain motivation over time [104, 106, 155].

In applied domains such as hiring, health, and education, studies emphasize the importance of aligning system automation with user expectations [69, 92, 114]. Employers often prioritize technical over interpersonal skills, revealing a potential disconnect with AI-driven assessments [77]. Across learning and review platforms, users benefit most when systems scaffold feedback exchanges, offer structure, and allow for agency-preserving participation [26, 67, 124]. Whether in peer feedback, educational critique, or idea evaluation, the tone, timing, and transparency of system responses significantly influence user motivation, reinforcing the value of dialogic and autonomy-supportive designs [38, 46, 153].

This study employs SDT as an evaluative framework to examine whether applicants feel supported in their needs for autonomy, competence, and relatedness during AI interviews. By testing different variants of response and feedback features, we explore how the interfaces affect applicants' sense of control, competence, and support throughout the process.

3 Study 1: Methodology

Study 1 aimed to examine applicants' perceptions and expectations of AI-driven asynchronous interviewers, particularly in the context of increasing access to LLM-based systems. We began by analyzing *Reddit* data, followed by a semi-structured interview study. To ensure a diverse range of perspectives, we selected discussions from several relevant subreddits. Based on the insights

[1] www.hirevue.com/
[2] https://quinncia.io/

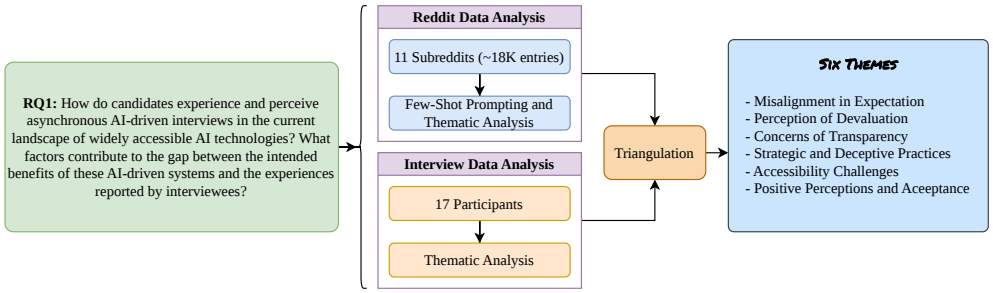


Fig. 1. Flow-diagram of Study-1 aiming to answer $\mathbf{RQ}_1$ through a qualitative analysis on *Reddit* discussion and interview data

gained, we conducted in-depth interviews to further explore emerging themes. The following section outlines the methodology of Study 1 in detail (shown in Figure 1).

3.1 *Reddit* Data Analysis

To explore the public discussion around AI-driven interviews and automated hiring, we analyzed *Reddit* data across communities where users discuss recruitment journeys, interview experiences, and algorithmic hiring systems. We initially identified a pool of subreddits based on prior literature [49, 53, 131] and curated lists from job-related websites [78, 113]. We then applied a structured filtering strategy to shortlist 11 subreddits that met the following criteria: (1) a minimum of 20,000 members to ensure visibility and community engagement, (2) thematic relevance to experience regarding AI-driven interviewers assessed through post content and flair usage, (3) sufficient discussion volume on AI-driven interviews (e.g., at least 50 relevant posts), and (4) public accessibility to allow ethical data collection. The final set of subreddits included *r/jobs, r/recruiting, r/recruitinghell, r/interviews, r/getemployed, r/antiwork, r/askHR, r/careerguidance, r/careers, r/careeradvice, and r/hiring*.

Using the PRAW API, we collected *Reddit* data from January 2023 to March 2025, applying a keyword filter derived from terminologies used in prior research to extract content related to AI-driven interviews such as “AI interviewer”, “Algorithmic hiring”, “automation in recruitment”; full list in Appendix A.1. This timeframe was selected to capture discussions following the introduction and widespread adoption of LLMs at the end of 2022. While some posts may reference interviews that occurred before 2023, their inclusion reflects how users chose to share or revisit those experiences during the period of increased public attention to LLMs. The initial dataset contained 25,154 entries. After manual review revealed many irrelevant entries, we refined the dataset using OpenAI’s GPT-4o model (temperature = 0.5) with few-shot prompting to identify entries describing experiences with AI-based asynchronous interviews, including justifications for relevance (shown in Appendix A.3) [150, 152]. The final dataset comprised 18,285 entries, including 9,707 unique users, 597 posts, 12,727 comments, and 4,961 replies. The average word counts were 225.4 (±305.1) for posts, 54.2 (±82.3) for comments, and 46.3 (±70.1) for replies. A Reddit post initiates a discussion thread, a comment is a response to the post, and a reply is a response to another comment within that thread.

Three human coders conducted the *Reddit* analysis using a coding reliability approach, following the procedures outlined by Braun and Clarke [20, 21]. We adopted a data-driven, grounded theory framework to allow codes and themes to emerge from the data rather than imposing predefined categories. During the coding process, they (1) reviewed the dataset filtered by the LLM to verify both the inclusion criteria and the rationale for filtering, and (2) conducted a open coding analysis.

For each subreddit, entries were first ordered by posting time from earliest to latest. We then drew an initial random sample of 100 entries to establish a baseline for code relevance and clarity. Our objective was to find themes from the posts and comments. Identifying the trend of discussion was out of scope of this paper. Thus, we decided to do random sampling from our dataset and intentionally avoided temporal data analysis. After this calibration step, thematic coding was conducted using randomly selected entries within each subreddit and continued iteratively until thematic saturation was reached, defined as the point at which no new codes emerged from additional data [127]. In total, we analyzed 1,893 entries (782 unique users, 197 posts, 1,696 comments and replies), and each coder judged over 93% of the LLM-selected data as relevant, supporting the appropriateness of the sampled content. Once the initial codes were identified, three coders met for several iterations to discuss their individually identified codes and through axial coding they grouped the related codes to form a refined and converged coding framework. Finally, a selective coding method was applied to identify the core themes of discussion of the *Reddit* communities on AI-driven. Based on the final list of core themes, three coders independently coded an additional set of 1534 entries (including posts, comments, and replies). To assess inter-coder reliability, we calculated Fleiss' Kappa coefficient, which yielded a value of 0.84, indicating strong agreement among coders [82]. The resulting themes represent the primary concerns and strategies shared by users across multiple subreddits. The distribution of themes across these subreddits is shown in Table 1.

Subreddit	Strategic Self-Presentation	Perceived Disrespect	Transparency	Expectation Misalignment	Positive Acceptance	Accessibility Challenges
r/jobs	1209 (32.66%)	715 (19.32%)	651 (17.59%)	467 (12.62%)	562 (15.18%)	97 (2.63%)
r/recruiting	299 (34.62%)	123 (14.26%)	227 (26.34%)	74 (8.60%)	138 (16.04%)	2 (<1.00%)
r/recruitinghell	696 (21.50%)	1221 (37.74%)	857 (26.48%)	178 (5.50%)	185 (5.73%)	98 (3.06%)
r/interviews	234 (37.95%)	143 (23.10%)	63 (10.15%)	108 (17.47%)	67 (10.85%)	2 (<1.00%)
r/getemployed	348 (31.51%)	215 (19.46%)	173 (15.71%)	223 (20.24%)	142 (12.82%)	3 (<1.00%)
r/antiwork	281 (10.61%)	1154 (43.51%)	490 (18.46%)	363 (13.70%)	240 (9.06%)	124 (4.67%)
r/AskHR	227 (16.69%)	363 (26.68%)	289 (21.19%)	189 (13.89%)	271 (19.90%)	23 (1.65%)
r/careerguidance	312 (13.31%)	317 (13.56%)	195 (8.34%)	739 (31.58%)	597 (25.50%)	181 (7.71%)
r/careers	38 (24.57%)	34 (21.68%)	20 (13.29%)	37 (23.80%)	26 (16.60%)	0 (<1.00%)
r/careeradvice	644 (34.38%)	337 (17.97%)	203 (10.83%)	386 (20.63%)	233 (12.44%)	70 (3.75%)
r/hiring	98 (25.60%)	38 (9.98%)	111 (29.03%)	58 (15.25%)	75 (19.65%)	2 (<1.00%)

Table 1. Distribution of the analyzed subreddit posts, comments and replies across different thematic areas.

3.2 Interview Data Analysis

Reddit analysis allowed us to develop our primary understanding of the challenges experienced by interviewees by analyzing a diverse, large-scale dataset. This approach was critical in developing a broad range of perspectives, especially from people who might be difficult to reach in formal study settings. However, analyzing posts and comments of Reddit users did not allow us to ask follow-up questions which is essential to uncover clarifying intent, context, underlying motivations, and underrepresented or absent nuances. To fill these gaps, we conducted a follow-up interview study.

Based on our *Reddit* analysis, we formulated a questionnaire for an in-depth interview study. Following the themes that emerged from the *Reddit* findings, we developed a semi-structured interview guide. The interview questions probed deeper to gather participants' experiences with AI-driven asynchronous interviews, focusing on areas such as perceived fairness, accessibility challenges, expectations versus reality, system transparency, data privacy concerns, and coping strategies used during the interview process (please refer Appendix A.2 for the interview questions).

ID	Gender	Ethnicity	Status	Major / Field	Neurotype	# AI Int.
P1	Male	Asian	Grad (PhD)	Computer Science	Neurotypical	3
P2	Female	White	Undergrad	Psychology	Neurodivergent (ADHD)	2
P3	Male	Black	Working	Business	Neurotypical	3
P4	Male	Asian	Grad (MSc)	Data Science	Neurotypical	4
P5	Non-binary	Hispanic	Working	UX Design	Neurotypical	2
P6	Female	Asian	Grad (PhD)	Education Technology	Neurotypical	3
P7	Male	White	Undergrad	Sociology	Neurotypical	2
P8	Male	Asian	Grad (MSc)	Computer Science	Neurotypical	6
P9	Male	Black	Grad (MSc)	Software Engineering	Neurotypical	3
P10	Female	Asian	Undergrad	Computer Science	Neurotypical	2
P11	Male	Multiracial	Grad (PhD)	Political Science	Neurotypical	3
P12	Male	White	Grad (MSc)	Mechanical Engineering	Neurotypical	3
P13	Non-binary	White	Grad (PhD)	Human-Centered Computing	Neurodivergent (ADHD)	2
P14	Female	White	Working	Communication Studies	Neurotypical	5
P15	Male	Black	Undergrad	Computer Science	Neurotypical	2
P16	Male	White	Undergrad	Computer Engineering	Neurotypical	2
P17	Male	Asian	Grad (PhD)	Software Engineering	Neurotypical	7

Table 2. Interview participants' demographics including gender, ethnicity, academic/work status, field, neurotype, and number of AI-driven interviews faced.

We invited participants from 12 U.S. universities using emails to student communities, alumni groups, and snowball sampling. Eligible participants were required to be at least 18 years old, proficient in English, and to have used an AI driven asynchronous interview system at least once. Before recruiting participants, we collected basic demographic information, including whether they identified as neurodivergent, as a significant portion of Reddit users with neurodiversity shared experiences related to AI interviews in these subreddits. We also asked preliminary questions about participants' AI interview experiences, including how many times they had participated and how recently their most recent interview took place. We initially received 51 responses and prioritized participants who had interacted with AI driven interviewers in 2024 or later. To maintain diversity and achieve thematic saturation without oversampling [127], participants were invited in batches of five, with recruitment continuing only when new or significant insights emerged. This process resulted in 17 completed interviews, with demographic details reported in Table 2. The interview study received IRB approval from the first authors' institution. Interviews were conducted online via Google Meet by the first and second authors and lasted 38 minutes on average. In total, we analyzed approximately 10.8 hours of interview recordings, resulting in about 90,000 words of transcribed text. Each participant received a $20 Amazon Gift Card.

We employed grounded theory (the same approach followed for analyzing the Reddit dataset) to analyze the interview data. The first and second authors independently conducted open coding following grounded theory principles [76, 138] for a subset of participants (N = 5), allowing conceptual codes to emerge inductively rather than imposing codes from the semi-structured interview guide. Through regular discussions and iterative refinements (axial coding), they consolidated these codes into a more organized coding structure [29, 75]. Finally, they identified the overall list of themes emerged from the interviews using selective coding. These themes were developed to capture the presence and range of issues discussed in the interviews, rather than to assess narrative depth or relative importance across data sources. Two coders independently coded the interviews of the remaining participants (N = 12) based on the identified list of themes. To check consistency, the inter-rater reliability was calculated using Cohen's Kappa, which yielded a score of 0.87, indicating strong agreement. To enhance the robustness of our findings, we applied triangulation [44], using

Main Theme	Sub-themes	Source
Misalignment Between Applicant Expectations	Expectation of Human-Like Presence (Anthropomorphism)	Both *Reddit* and Interview
	Discrepancy Between Promotional Language and Actual Functionality	Both *Reddit* and Interview
	Expectation of LLM-like Capabilities	Both *Reddit* and Interview
	Difficulty for Non-Native Speakers Due to Accent Handling	Interview Only
Perception of Disrespect and Devaluation	Disrespect Stemming from One-Way, Non-Interactive Format	Both *Reddit* and Interview
	Dissatisfaction Directed at Hiring Process, Not AI Itself	Interview Only
	Superficial Encouragement Without Genuine Responsiveness	Both *Reddit* and Interview
Concerns Regarding Transparency	Insufficient Communication from Hiring Organizations (External Transparency)	Interview Only
	Uncertainty About System Functioning and Evaluation Criteria (Internal Transparency)	Both *Reddit* and Interview
	Data Privacy and Security Concerns	Both *Reddit* and Interview
	Perceived Exploitation of Applicant Data	*Reddit* Only
Strategic Representation and Deceptive Practices	Strategic Performance and Impression Management	Both *Reddit* and Interview
	Leveraging External Tools to Bypass System Constraints	Both *Reddit* and Interview
	Moral Conflict and Justification Under Pressure	Both *Reddit* and Interview
Accessibility Challenges for Neurodivergent Candidates	Risk of Misinterpreting Neurodivergent Behaviors	Both *Reddit* and Interview
	Reluctance to Disclose and Seek Accommodations	Interview Only
	Emotional Strain from Lack of Human Feedback	Both *Reddit* and Interview
Positive Perception and Acceptance	Convenience and Flexibility	Both *Reddit* and Interview
	Perceived Standardization and Fairness	Both *Reddit* and Interview
	Pragmatic Adaptation to an Inevitable System	Both *Reddit* and Interview

Table 3. Thematic Analysis of Study 1, presenting main themes, their sub-themes, and associated data source(s). Color codes are used to distinguish each main theme, and the same colors are applied to their corresponding sub-themes to maintain visual consistency when they are discussed later.

interview data to clarify and contextualize themes identified in the *Reddit* analysis rather than to introduce new dominant categories.

3.3 Consolidation of Reddit and Interview Themes

We used a convergent triangulation approach to thematically integrate the themes that emerged from the *Reddit* and interview datasets, consistent with prior qualitative research [25, 55, 111, 132]. We iteratively examined the alignment between themes identified in the Reddit data and those observed in the interview data, without treating either source as a primary standard. This process allowed us to identify areas of convergence (themes present in both sources), complementarity (themes that elaborated on each other), and divergence (themes unique to one source). The themes supported by both data sets were merged and those unique to one data set were retained as such, with their origin explicitly indicated in Table 3. Throughout the analysis, we maintained source tags (*Reddit* vs. interview) for all coded excerpts, which allowed us to trace how evidence from each dataset contributed to each sub-theme.

4 Study 1: Result

In the following section, we present the key themes identified from our *Reddit* and interview data. Interview participants are labeled as 'P', and Redditors as 'R', with *Reddit* comments numbered sequentially for clarity. Most of the Redditors quoted in this section are unique users (25 out of 28). For ethical purpose, we have paraphrased all the quotes we collected from *Reddit* without removing their essence and meaning. Table 3 presents all the main themes identified in our analysis, along with their corresponding subthemes and associated data sources (interviews, *Reddit*, or both).

Theme 1: Misalignment Between Applicant Expectations

Participant/ Redditor ID	Quotes
P2	As they said, I thought it'd be smart, like actually "do" something. But nope, just a blank screen and questions. Not even a personalized feedback.
P3	The whole thing felt like a marketing pitch gone wrong. They sold it like it was high-tech, but it didn't even try to respond or adapt.
P7	If they just said it was pre-recorded, fine. But calling it 'AI' made me expect a conversation. There's a responsibility in how you frame tools.
R1	...Not sure if anything has changed. I worked with P&G years ago and their AI interview tool looked fancy but gave zero context. I expected today's tools to interact at least a little, it didn't...
R2	...They said 'AI-powered,' I thought cool, like ChatGPT or something. Turns out it was just a glorified pre-recorded form. Disappointing.
R3	If they just told us upfront it was static, I'd manage my expectations. But calling it 'conversational' set me up for a tech letdown. I mean, just be clear next time.

Table 4. Interview participants and Redditors' paraphrased quotes highlighting expectation mismatches in AI-driven interview experiences.

4.1 Theme 1: Misalignment Between Applicant Expectations and AI-Driven Interview Processes

The mismatch of expectations emerged as one of the most prominent themes in our study. A substantial number of *Reddit* comments and interview responses (number of *Reddit* entries or NR= 66, number of interview participants or NI=8) pointed to elements of anthropomorphism, such as remarks like "the system does not feel like a human" or "I cannot see the nonverbal cues during the interview," which have been explored in prior research [15, 30, 68]. However, our findings revealed

two unique types of expectation discrepancies: (1) applicants observed a noticeable gap between the way interview tools were described and how they actually functioned (as pointed out by P3 and R2), and (2) applicants anticipated more advanced, conversational, and interactive systems that resembled contemporary AI technologies rather than the impersonal and static tools they encountered (shown in Table 4).

Several interview participants (NI=5) noted that the instruction emails sent by the hiring organizations used phrases such as "AI-based," "with our AI tool," or "our conversational platform," (an issue raised by P7 and R3) which led them to expect a technologically advanced system. In some cases, participants were provided with a practice link intended to help them become familiar with the platform. However, when they accessed the actual interview, the system appeared noticeably different- both in design and interaction style and often felt more like a static asynchronous video recording tool than the advanced technology they had anticipated (echoing concerns from P2 and R1). These findings reiterate that terminologies can create inflated expectations, and when unmet, lead to disappointment and distrust of the AI-systems (a reaction mirrored by R3 and P3) [83]. As we heard from P2 said,

> *"[...] they (hiring organization) said their AI recruiter would take the interview, but honestly, the system didn't live up to that. The only AI thing about it was some avatar, and even that didn't do much. You just answer the questions, and that's it. I was expecting something more conversational, like how ChatGPT talks back, you know? After a while, it became hard to keep up the energy and interest. There was no back and forth, no real interaction, nothing that felt intelligent. It just started with a fixed set of questions, and no matter what I said, some generic text would pop up on the corner. It just blindly followed the list. I was hoping for something that could actually adapt to what I was saying."*

P2 added that the evaluation might have happened in the background by an AI-based algorithm but knowing it beforehand would have been better for them to prepare mentally what to expect on the spot. We noticed applicants (NR=31, NI=6) increasingly expected a more conversation-like experience similar to their interactions with LLM-driven chatbots (e.g., ChatGPT, Gemini etc) as evident in our findings where R2 explicitly compared their expectations to "ChatGPT-like" interaction and R4 was expecting the system would build on their answers like an LLM does. R5 commented where the OP (original poster) shared an experience with an AI-driven interviewer:

> *"[...] there needs to be better communication or use of this tool. As the job market has become tighter, people already feel devalued when they are interviewed by an automated system. And it gets worse when it is not even up to the standard. Basically, you are just talking to a box at best and being profiled at worst."- R5*

Apart from the lack of conversational features, non-native English-speaking participants (NI=3) shared their struggles with systems that poorly transcribed their responses. Prior experience with LLMs shaped expectations of comparable support from the interviewer. However, the experience was otherwise, as the transcription errors affected how their answers were represented, leading to frustration and a sense of being misjudged. When asked how they knew the system was not using an LLM to capture their response, P1 answered:"

> *"I know based on past experiences with Siri and Alexa. These systems tend to struggle with non-major English accents. That's why I try to enunciate with an American accent when speaking to these tools. But I saw that ChatGPT can capture my sentences even when I am speaking with my natural accent. That's how I know these tools (AI hiring tools) are not using LLM features."- P1*

Another participant, P8, added that masking an accent is itself a stressful task. But knowing that their accent could be the biggest barrier for them to secure their desired job made the experience even more frustrating. They shared:

> *"You know you are already at a disadvantage because English is not your first language, and I admit I should be more fluent in speaking. But when you are in an interview and the experience makes you feel excluded, even when the technology exists to do better, it just makes you feel small."- P8*

4.2 Theme 2: Perceptions of Devaluation in One-Sided or Non-Interactive Interview Formats

Theme 2: Perceptions of Disrespect and Devaluation

Participant/ Redditor ID	Quotes
P8	If it's gonna replace a person, it needs to at least feel present. This felt like a surveillance camera with a mic
P11	...It wasn't even the tech that bothered me, it was the complete lack of acknowledgment. Like, I said all that just to get a generic 'thank you for your submission'?...
P10	It's not even about being anti-AI. I'd take a smart bot over a rude recruiter any day. But this? This was just me performing to silence. No response, no presence, just me and my doubts.
P15	It so feels like a cog in the wheel. You are definitely not valued as a human, just another datapoint in their pipeline.
R6	...The system said 'you're doing great' after every question. Sweet, but who decided that? It felt less like reassurance and more like automated flattery...
R7	...If this is the future of hiring, count me out. I don't need a robot to smile at me, but at least pretend like my response mattered...
R8	...There is no way in freaking hell I will ever do another one-way video interview. It's awkward, unnatural, and makes my anxiety spike through the roof. At least with a real person, I can read the room....

Table 5. Interview participants and Redditors' paraphrased quotes highlighting perceptions of disrespect and devaluation in one-sided interview formats.

A recurring theme we found was the applicants' perception of disrespect and devaluation within the one-sided, asynchronous AI-driven interview format. Prior research has similarly documented resistance toward AI technologies when they are perceived as replacements for human roles [137, 148]. Although many comments on *Reddit* (NR=42) expressed strong aversion toward AI-driven interviewer tools, we observed a number of participants from the interview study are not entirely opposed to AI-driven interviewers themselves. Rather, their dissatisfaction is more directed toward the broader hiring process and the rigid, non-interactive interview format as also described by P10 calling it "performing to silence" (shown in Table 5). While various design considerations have been proposed to improve the asynchronous interview experience, this gap in user acceptance and perceived fairness remains unresolved. As R9 commented on *Reddit*,

> *"I have done numerous interviews with AI.[..] so it's not that people are resisting the use of AI in the hiring process, they're just resisting the disrespectful model of the one-way video interview."- R9*

This tension highlights a deeper issue: advancements in AI, have not translated into a more human-centered interview experience. Despite improvements in interface design and system

responsiveness, applicants continue to feel unseen and devalued suggesting that the core problem lies not in technical capability (as reported by P10 and P11), but in the lack of reciprocal, relational interaction. P15 mentioned that the recruitment tool he faced used LLM as per the description in their website. The overall system felt very advanced, and the interface was highly intuitive, but they still felt devalued by the one-sided nature of the interview:

> *"There might've been an LLM behind the screen, but to me, it still felt like no one was listening."- P15*

Another participant, P16, shared a similar experience, noting that the system was smooth and the supportive text pop-ups, such as brief prompts like *"Take your time"* or *"You are doing great"* helped create a more comfortable environment. However, they still felt unacknowledged during the interaction, as their responses received no apparent recognition which reinforced the sense of a one-sided exchange (as R6 described it as a hollow reassurance). This impression was further reinforced when the final feedback appeared overly generic, suggesting that their answers had not been individually considered.

> *"[...] It (the final feedback) all looked polished, but it felt like the outcome was decided before I even spoke, more like an automated summary than a response to what I actually said."- P16*

4.3 Theme 3: Concerns Regarding Transparency and Algorithmic Decision-Making

Theme 3: Concerns Regarding Transparency

Participant/ Redditor ID	Quotes
P1	I wasn't asking them to explain the entire algorithm. But just a simple breakdown of what's being evaluated, like whether it's just verbal response or also non-verbal cues would've helped me prepare more confidently.
P4	Even after completing the process, I wasn't sure if a human ever saw my video. It's hard to reflect or improve when you don't even know who or what was reviewing your performance.
P9	I think transparency could actually improve trust. A little more clarity around what's analyzed and how data is handled would go a long way toward making the experience feel more ethical and respectful.
R10	...They call it 'interview automation' but it's just secret surveillance wrapped in startup buzzwords. If you're gonna judge me for eye contact, at least say so...
R11	...Not HireVue, but I did one where I had to answer timed questions on camera. I kept thinking 'If I pause to think, will it look like I'm unprepared?' My anxiety made me rush, and I hate how that probably affected how I came off...
R12	...Like others have said, interviews are a two way street. At the very least, tell me what you're scoring or looking for. If it's all automated, that's fine but let me know. Otherwise, it just feels manipulative.
R13	...Probably free data for their AI, not a real interview. They just want a wide range of responses to train their algorithm. No real interest in hiring....

Table 6. Interview participants and Redditors' paraphrased quotes highlighting concerns regarding transparency and algorithmic decision-making.

Transparency has been a major topic of discussion for black box systems. There are many concerns reported about the AI-driven interviewer in previous studies, which we also observed in our study [7, 9, 130]. However, our findings focused more on the transparency related to both

external (communication from hiring organizations) and internal (functioning of the AI-driven systems) aspects of the interview journey. We learned from interviewees that applicants are often less informed or not informed at all about critical interview-related details. Although it is understandable that hiring organizations may not reveal everything about their hiring process, as per the participants, organizations failed to provide "necessary details" to make the experience better as echoed by P1, who wished for even a basic outline of what would be evaluated, including non-verbal cues (shown in Table 6).

Participants described gaps in both external and internal transparency that shaped their experience of the interview. Several interview participants (NI=6) noted that, in some cases, they were not given clear guidelines on how to approach the interview. This pointed to a gap in *external transparency*, as organizations provided only general instructions about logistics, such as checking the internet connection or using a computer, without clarifying the interview format or what would be evaluated. At the same time, participants (NR=39, NI=8) reported a lack of *internal transparency* about how the AI-driven systems operated. There was little to no information about what to expect from the system itself. Although the tools appeared advanced in terms of interface design, applicants were unsure whether behaviors such as facial expressions, eye contact, or body language would be assessed. R10 described this uncertainty as "secret surveillance wrapped in startup buzzwords." This confusion was also reflected in P17's experience, who shared that before their first AI-driven interview, the organization sent an email that referred only to a "standard interview procedure," which gave basic setup tips but offered no explanation of how the system functioned or what it would assess.

> *"I was really confused. The good part was that the system looked impressive, but since it was my first attempt, and there were no proper guidelines, it felt like I was taking a test without being told the rules. I had no idea if I was being recorded for an actual evaluation or just for show. Was the system monitoring my facial expressions? Should I keep eye contact all the time or look away occasionally? Should I smile while answering? Should I wear formal attire? All these questions continued to run through my mind. And that kind of guessing game really messes with your confidence."- P17*

Another major concern revealed in our findings was the data privacy issues. Participants (NR=78, NI=7) voiced significant privacy concerns regarding their recorded responses. While hiring organizations captured their video interviews, they were not clearly informed about how these recordings would be stored, who would have access to those recordings, or how long the data would be retained. This lack of transparency raised fears about potential misuse, unauthorized sharing, or data breaches. R14 commented on a class action lawsuit in which a major pharmaceutical company is sued for secretly using AI technology to analyze facial expressions during interviews and assigning candidates an 'employability score' without informing them [32],

> *"After that whole CVS and HireVue mess where they secretly tracked people's facial expressions and scored them without telling anyone, it is no wonder nobody trusts AI interviewer agents anymore."- R14*

This fear of data exploitation was also shaped by other experiences. Redditors (NR=13) mentioned feeling disrespected by the one-sided format of AI-driven interviews, facing rejection without explanation, and often being ghosted after completing the process [97]. These experiences led some to conclude that the organizations were using the interview primarily as a means to harvest data rather than to truly consider them for the position (as R13 described "free data for their AI"). As shared by one of the Redditors, R15:

> *"It felt less like a job interview and more like unpaid data entry for their algorithm."- R15*

Our study revealed concerns about the lack of transparency in how AI systems make decisions, particularly in evaluating interpersonal behaviors. One common example was eye contact. Participants were told to maintain eye contact during interviews but were not informed how the system interpreted it. Many looked away briefly to think and later got worried that this might be interpreted as disengagement. Moreover, it was unclear whether a human would later review their responses or if the evaluation was entirely automated (as mentioned by P4). This uncertainty left participants unsure whether to use specific keywords to appease an algorithm or speak naturally to connect with a possible human reviewer.

4.4 Theme 4: Strategic Self-Presentation and Deceptive Practices in AI Interview Responses

Theme 4: Strategic Self-Presentation and Deceptive Practices

Participant/ Redditor ID	Quotes
P3	Of course you can game the system. You can game anything once you figure out how it scores. AI interviews are just another test. You learn the rules, then play to win.
P12	I filtered my answers. Not because I wanted to, but because I felt like the system couldn't handle nuance.
R16	...AI interview has generic questions like, 'Tell me about a time you worked on a team.' It's less about being honest and more about hitting the right phrases. You start answering like you're writing SEO copy...
R17	...So how exactly does an AI interview do that? You just trick the system by giving overly structured, polished answers. It's not about who you are, it's about how well you play the game. It's not even real...
R18	...I tried it out to see what response I'd get. And it felt like the whole thing encouraged me to rehearse, not reflect. The moment I tried to speak naturally, I worried I'd be penalized for going off-script....
R19	...Do you mind if I use your video but instead of your voice, I dub in an AI-generated voice and see if I get more hits? Just testing the system's biases...

Table 7. Interview participants and Redditors' paraphrased quotes highlighting concerns regarding strategic self-presentation and deceptive practices in AI interview responses.

From *Reddit*, we collected numerous comments (NR=167) that largely focused on how to navigate AI-driven interviews and the deceptive practices that some applicants use to get through them (shown in Table 7). Since our selected subreddits were about tips and tricks to do well in the interviews, many Redditors (e.g., R16, R17, R18) shared their experience to help others but at the same time we found people mentioning about deceptive strategies which they adopted often in time of desperation.

As part of their self-presentation strategies, several interview participants and *Reddit* users (NR=41, NI=9) described intentionally using certain keywords during the interview. This aligns with prior findings noting that Applicant Tracking Systems (ATS) often scan for such keywords as part of automated evaluation processes [24]. Some of the users (NR=18, NI=5) reported exaggerating positivity, energy, or expressions (e.g., forced smiles or sustained eye contact) in response to assumed non-verbal tracking by the system, a tactic known as deceptive impression management (IM) by faking impression also in prior research [134]. Some shared to memorize pre-written responses with correct pronunciation to common questions to appear more polished to create positive impressions.

One of the participants in our interview study (P14) described preparing for the interview "like a hackathon" emphasizing the strategic nature of performance under constrained interaction.

> *"I planned it like a hackathon. Calm face, solid voice, keywords on cue. Can't risk surprises with a system that can't ask follow-ups."- P14*

This kind of performative action was seen to be exhausting. For example, R20 humorously captured this pressure by saying:

> *"At this point, I'm not sure if I'm applying for a job or auditioning for an Oscar. Smiled like a maniac, nailed the keywords, and prayed that the AI liked my lighting setup. Still got ghosted"- R20*

Reddit users (NR=26) often described using external aids to game up the interview process. Many were frustrated by the lack of non-verbal cues and the limit on re-recordings, which made it hard to refine their answers. In some cases, they had no retry option and had to do it in a single attempt. Therefore, they turned to LLMs to craft polished responses on the first try and even used AI voice bots to deliver them as described in a comment "..just let AI talk to AI". While we do not endorse such practices in any circumstances, they point to a broader issue: When systems feel rigid or opaque, applicants may look for workarounds. Some users expressed guilt, but justified their actions due to high stakes, perceived unfairness, or a belief that the system was already biased. P3 shared that going through a one-sided interview with templated, inflexible response formats was particularly challenging for them. They added that, in such situations, they could understand why someone might feel compelled to bend the rules or seek outside help just to get through the process.

> *"I've heard of people using ChatGPT or notes to get through these interviews. I never did it myself, but I can understand why someone would, especially if they felt they weren't being given a fair shot. Sometimes it's not about cheating, it's about surviving a system that already feels stacked against you."- P3*

4.5 Theme 5: Accessibility Challenges for Neurodivergent Candidates in AI-driven Interviews

While prior research has examined accessibility and design concerns through structured interviews [8], our analysis of unsolicited narratives from public forums such as *Reddit* suggests additional areas that may warrant attention. These include possible misinterpretations of neurodivergent behaviors, emotional challenges associated with systems lacking human feedback, and the ways in which some users may adapt strategically to such environments. While some users (NR=9, NI=2) shared coping strategies (see Table 8), others reported consistently negative outcomes, particularly when AI systems appeared to evaluate non-verbal behaviors that may differ from neurotypical norms. One Redditor, R24, reflected:

> *"I'm autistic and have never even landed a second interview after a Workday screening. I looked into it and found that the AI evaluates facial expressions, eye contact, and body language to predict job performance. It feels completely discriminatory."- R24*

Despite companies offering adjustments, this support often depends on candidates self-identifying and requesting accommodations [13], which many are reluctant to do. This hesitation often leads neurodivergent applicants to withhold their condition and instead find alternative, less direct ways to manage the process. As P13 shared:

> *"I no longer disclose my condition because the accommodations offered have been ineffective or have ended up working against me. Instead, I ask to spread long interview rounds over*

Theme 5: Accessibility Challenges for Neurodivergent Candidates

Participant/ Redditor ID	Quotes
P2	I really tried to stay focused, but staring at the camera with no feedback just made my brain wander. With ADHD, I rely a lot on cues, tone, facial cues, little reactions - to stay anchored. This was just silence. By the third question, I was already mentally checked out. **(ADHD)**
P13	I spent more time prepping how to act than how to answer. Should I pause more? Smile? Look at the lens the whole time? ADHD makes me second-guess everything in that kind of setting. And because there was no feedback, I kept overthinking it even after it ended. **(ADHD)**
R21	...I had an interview where the company asked me to keep my eyes on the camera the whole time. I couldn't focus on the question and think clearly while also staring at the lens. I kept wondering if looking away for a second would cost me... **(Autism)**
R22	...What's annoying is that when I didn't use AI to help write my answers, I got ghosted. But when I used it to 'polish' my wording, I got callbacks. Makes you wonder how much authenticity actually matters or if it just rewards people who mask better... **(Autism)**
R23	...Same thing happened to me... when they called, I disclosed that I have anxiety and ADHD. After that, I never heard back. Automated systems don't care about that stuff.... **(ADHD)**

Table 8. Interview participants and Redditors' paraphrased quotes highlighting concerns regarding accessibility challenges for neurodivergent candidates in AI-driven interviews.

multiple days, but frame it as a scheduling issue rather than a disability-related need."- P13

Neurodivergent individuals often navigate social interactions by interpreting subtle contextual cues; intonation, pacing, facial feedback that help them adjust and respond appropriately [157]. In digital systems that lack these human elements, especially AI-driven interview platforms or automated assessment tools, this interpretive space collapses. The absence of reassurance or clarification can lead to heightened anxiety, self-doubt, and a persistent feeling of being misjudged or misunderstood. Rather than adapting to the system, users may find themselves overcompensating or disengaging entirely. These experiences highlight the emotional friction that arises when interfaces are designed without responsiveness or recognition of differences. One Redditor, R25, who is autistic, shared that they did not receive additional accommodations because the hiring organization did not have such a provision. They proceeded with the interview anyway, but found it to be a rather unpleasant experience:

"It's like talking into a mirror that doesn't reflect anything back. I can't tell if I'm being understood, or if I've made a mistake. So I rehearse every word, not because it helps, but because I'm scared [that] the system will see me as wrong just for being me."- R25

4.6 Theme 6: Positive Perceptions and Acceptance of AI-Driven Interview Systems

Despite various flaws and limitations, both *Reddit* comments and interview participants (NR=61, NI=10) highlighted several positive aspects of AI-driven interview systems (shown in Table 9). Many Redditors appreciated the speed, scheduling flexibility, and automation these systems offer. Such features were seen to alleviate some of the logistical burdens typically associated with traditional interviews. P17 described the system as "convenient" when applying to multiple positions. In particular, one-way or pre-recorded formats enabled some users to prepare in a low-pressure environment and better manage social anxiety. As one Redditor, R28, shared:

Theme 6: Positive Perception and Acceptance

Participant/ Redditor ID	Quotes
P4	Honestly, I liked how structured it was. No small talk, no guesswork, just give your answer and move on.
P17	The process was fair in the sense that everyone got the same setup. No interviewer bias, no weird vibes, just consistency.
R25	...Was it weird talking to my webcam? Yeah. But I did it at 2 AM in pajamas with my dog next to me. Try doing that in a regular interview...
R26	...Say what you want about robots, but at least they don't care if you forgot to iron your shirt or have a weird laugh. Everyone gets the same setup, and honestly, that feels fairer sometimes..
R27	...Look, AI interviews aren't perfect, but neither is traffic or 9 AM calls. The tech will get better, but meanwhile, I'd rather learn how to ride the wave than wait for it to crash...

Table 9. Interview participants and Redditors' paraphrased quotes highlighting positive perception and acceptance of AI-driven interview systems.

> *"I prefer recording answers alone rather than speaking to a panel. It was better than being stared at by four people on Zoom."- R28*

Some Redditors viewed AI-based interviews as more standardized or impartial, suggesting that they reduce certain human biases and promote equal treatment. Similar views emerged in our interviews. P9 noted that while the system's uniformity ensures everyone faces the same environment, this can be both its strength and its limitation:

Others took a pragmatic stance, recognizing that AI-driven interviews, although imperfect, have become a routine part of hiring. Rather than dwelling on their flaws, participants emphasized adapting and moving forward, noting that earlier formats had their own limitations as well.

P9 further remarked that when system shortcomings are evident, hiring organizations have the responsibility to address them. A well-designed tool is not enough on its own; it must be paired with thoughtful implementation:

> *"The issue arises when the system is mediocre and the hiring organization is also ignorant about their responsibilities. Undoubtedly, they can be more sincere." – P9*

Some users also reflected on how they learned to adapt to the system by understanding how AI evaluates responses. They argued that while the process may be uncomfortable, resisting it is less productive than learning how to navigate it, especially as it becomes more common in hiring.

4.7 Summary

Our findings reveal a disconnection between applicants' expectations and their actual experiences with AI driven interviews. Many expectations were shaped by the language used by hiring organizations and the growing public familiarity with LLMs (e.g., ChatGPT, Gemini), which led some to assume that the systems would be highly responsive and personalized. Participants expressed disappointment when the interviews appeared conversational but lacked personalized interaction or adaptation. Several applicants described using strategies to trick or bypass the system, often justifying their actions by pointing to its rigidity and perceived lack of fairness. Neurodivergent individuals reported experiences of exclusion and shared specific workarounds they developed to cope with the process. While many described the experience as dehumanizing, some acknowledged

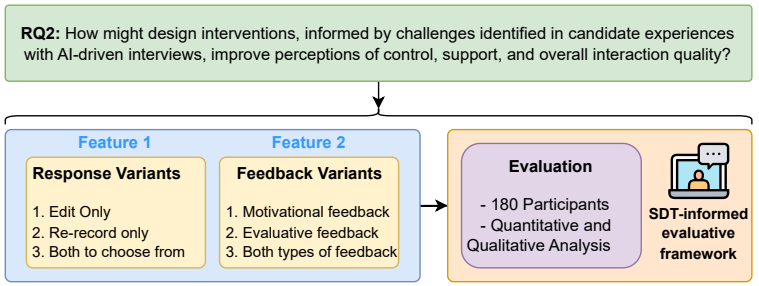


Fig. 2. Flow-diagram of Study-2 aiming to answer $\mathbf{RQ}_2$ through different variants of two features separately; response type and feedback type followed by a mixed-method analysis

benefits such as consistency and efficiency. Still, acceptance was often marked by resignation rather than trust, shaped by the absence of better alternatives.

5 Study 2: Methodology

Building on the findings of Study 1, we conducted a follow-up study to explore design opportunities for improving applicant experience. Study 1 showed that while participants acknowledged the benefits of AI in recruitment, many participants were not entirely positive about AI-driven interviews. After becoming familiar with LLMs, their expectations had shifted, and they wanted interviews to feel more personalized and human-centered. They often felt devalued, seeing this as a result of organizational neglect or the tool's limitations, and wanted more acknowledgment during their interactions with the AI-driven interview system. Guided by these insights, we considered features for an asynchronous AI-driven interviewer to facilitate a stronger sense of personalization and value. We refer to each design consideration as a *feature* and its specific interface implementation as a *variant*. The following section explains our design considerations in detail, followed by the interface implementations and the data analysis approach (see Figure 2).

5.1 Design Considerations and Motivation

Our initial study showed that many applicants felt disrespected by the one-sided nature of AI interviews, citing a lack of acknowledgment and control over how they could present themselves. These concerns reflected broader issues of motivation, fairness, and self-representation. To address them, we focused on two key directions that consistently emerged across interviews and *Reddit* discussions: enhancing applicants' sense of agency and their sense of belonging. Based on these insights, we established the following design considerations (DCs):

5.1.1 ***DC1:*** *Give interviewees more agency to ensure their responses are interpreted as intended.*

Interview participants (e.g., P1, P2, P3, P7, P10, P12, P14, P17) and Redditors (e.g., R3, R5, R8, R14, R19) described feeling devalued during AI-driven interviews because they could not tell how the system interpreted their answers. Many noted that organizations gave little explanation or transparency about how responses were assessed, which made them feel disregarded. This uncertainty led them to second-guess their words, remove nuance, or mimic what they thought the system wanted, which reduced the authenticity of their self-presentation. Several reported over-rehearsing or changing their tone and accent out of fear of being misunderstood. With the human element removed from these high-stakes evaluations, they felt organizations had a greater obligation to ensure fairness and clarity. As one participant said, *"I wasn't asking them to explain the entire algorithm. But just a*

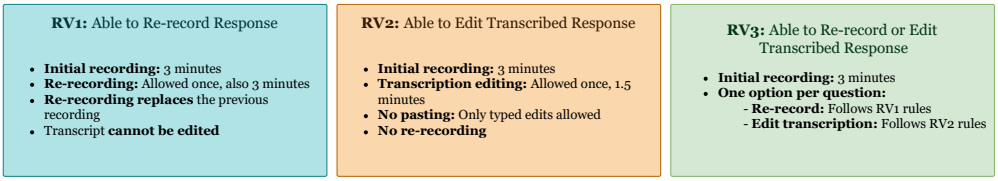


Fig. 3. Design details of different variants of response interface: RV1 only lets re-record, RV2 only lets edit and RV3 provides both options to choose from

simple breakdown of what's being evaluated… would've helped me prepare more confidently" -P1. These challenges show why giving interviewees more agency is vital. Supporting their sense of autonomy can enable authentic self-expression, reduce anxiety about misinterpretation, and help them present their abilities with confidence.

5.1.2 ***DC2:*** *Give interviewees personalized feedback to help them feel acknowledged and valued.* Interview participants (e.g., P2, P8, P10, P11, P15, P16) and Redditors (e.g., R2, R6, R9, R13, R21) repeatedly noted that AI-driven interviews made them feel invisible because they received no acknowledgment or personalized response to their efforts. Many described putting substantial thought into their answers only to get generic confirmation messages, which made the process feel mechanical and dismissive. This lack of recognition eroded their sense of belonging and made them feel like "just another entry in a dataset." One participant reflected, *"...It wasn't even the tech that bothered me, it was the complete lack of acknowledgment..." -P11.* A Redditor echoed, *"It felt like yelling into a void. If they can analyze me, they can at least tell me something back" -R6.* These reactions show why personalized feedback is essential: it would counteract feelings of devaluation, signal that their individual contributions matter, and help restore a sense of human connection in the interview process.

5.2 Interface Design and Implementation

For $\mathbf{DC_1}$, we envisioned several design choices and decided to include a feature that allows applicants to re-record their responses. Our goal was not only to display their recorded response as transcribed text but also to let them correct unintended errors. We designed three **response variants (RVs)** for this interface. The first variant, *Re-record only*, lets users record a new response if they find the transcription significantly inaccurate. Although several existing AI interview systems allow re-recording, we included this variant mainly as a baseline for comparison. The second variant, *Edit only*, allows users to revise the transcribed text within a limited time window. This design was informed by participants who shared that they sometimes wanted to make small changes or add details but were hesitant to start over from scratch. The third variant, *Both*, allows users to choose between re-recording and editing. We hypothesized that providing both options would enhance their sense of agency and reduce stress [105]. The design of these variants is shown in Figure 3.

For $\mathbf{DC_2}$, our objective was to provide some form of personalized messages or responses to users so that they can feel valued. One way to achieve that goal is through personalized feedback. Prior research suggests that such feedback can enhance user engagement and trust in AI systems [41, 57]. After careful consideration, we designed another interface with three types of **feedback variants (FVs)** to fulfill $\mathbf{DC_2}$. (1) motivational feedback, aimed at acknowledging participants' effort and maintaining engagement [59]; (2) evaluative feedback, offering real-time suggestions [95]; and (3) combined feedback, which included both types of feedback (see Appendix A.5). Motivational feedback was designed to reinforce reflection, and emotional safety, deliberately avoiding critique

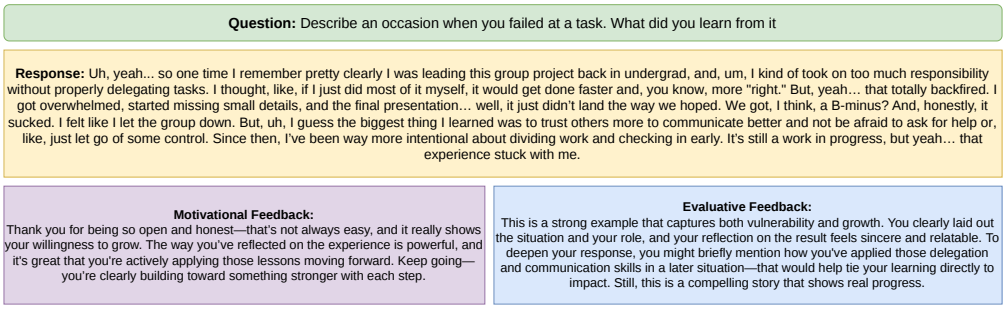


Fig. 4. An example of motivational and evaluative feedback for a behavioral interview response to *"Describe an occasion when you failed at a task. What did you learn from it?"*. The motivational feedback emphasizes encouragement and personal growth, while the evaluative feedback provides constructive suggestions grounded in the STAR (Situation, Task, Action, Result) framework.

or correction. Evaluative feedback, in contrast, offered warm but targeted suggestions in-between the questions to help users refine their responses based on the STAR framework (Situation, Task, Action, Result) which is a common rubric to evaluate behavioral response [3] (shown in Figure 4). We avoided using any numerical scoring as we did not want to cause performance anxiety or comparison pressure. In the combined condition, both types were shown, with motivational feedback presented first to align with positive framing principles [149].

For our study, we developed a web-based interface that simulated an AI-driven interviewer system using Streamlit[3], customized with CSS to control the layout and visual design. To provide participants with a realistic interview experience, we instructed them to treat the session as a mock interview. Audio recordings were transcribed via OpenAI's Whisper API[4]. To support our experimental design, we built separate interfaces for RVs and FVs. As these two features had three variants, we ended up building in total of six different variants. In the RVs, participants could either re-record or edit their response. But in the FVs, participants were not allowed to revise their responses. Instead, they viewed the transcription and received personalized feedback generated by GPT-4o through few-shot prompting (see Appendix A.5). We recorded all transcriptions, feedback, and system activity logs for analysis. The interface was designed to be minimal and intuitive, supporting ease of use. The application was hosted online to enable remote participation (shown in Figure 5).

5.3 User Study and Data Analysis

5.3.1 Theoretical Framework and Measures.

To evaluate and compare the efficacy of each variant, we used *Self-Determination Theory (SDT)* as a guiding framework, following prior HCI and CSCW studies that applied SDT to design and evaluate interactive systems [37, 57, 123, 154]. We chose SDT because our design considerations align with its three psychological needs: autonomy, competence, and relatedness. $\mathbf{DC_1}$ (response variants or RVs) was expected to support autonomy by allowing users to control how their responses were captured, and competence by helping them express themselves more effectively. $\mathbf{DC_2}$ (feedback variants or FVs) was intended to support relatedness by helping users feel acknowledged and emotionally supported, and to support competence by boosting their confidence through constructive guidance.

[3] https://streamlit.io/
[4] https://openai.com/index/whisper/

Fig. 5. Snapshot of our system: Shows the response variant-3 or RV3, where the user can see both options (edit or re-record) after recording their response (above), and the feedback variant-1 or FV1, where the user can only get motivational feedback (below)

We developed custom self-report items inspired by SDT-informed practices in HCI research. For the RVs, we asked participants if they felt a greater sense of control (autonomy), if they could express themselves more effectively (competence), if they felt confident in their self-presentation (competence), and if they found the feature useful (a pragmatic usability measure). For the FVs feature, we asked if they found the feedback useful (usability), if it increased their confidence in their competence (competence), if they felt acknowledged (relatedness), and if it enhanced their sense of engagement (motivation). While these items were not taken from a standardized SDT scale, they follow established SDT-based HCI practices that use task-specific measures aligned with autonomy, competence, and relatedness [61, 154].

5.3.2 Participants.

We recruited 180 participants from 19 U.S. universities through student community emails, alumni

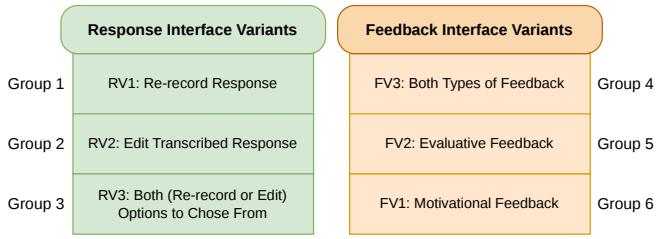


Fig. 6. Group-wise allocation of different variants of response and feedback interfaces. (RVs = Response Variants, FVs = Feedback Variants)

networks, and snowball sampling. Eligibility criteria included being at least 18 years old, fluent in English, and having prior experience with AI-driven asynchronous interviews. Participants from Study 1 were excluded to minimize bias. Each participant received $10 USD after completing the study. The study received approval from the authors' Institutional Review Board (IRB) office.

Participants ranged in age from 18 to 36 years. Of the 180, 102 identified as male, 64 as female, 10 as non-binary or another gender, and 4 preferred not to say. Ethnic backgrounds included Asian (74), White (54), Black or African American (34), Multiracial (10), and Other (8). Most participants were graduate students (92), followed by undergraduates (46) and employed individuals. All participants had prior experience with AI-driven interviews: 72 had used them once, 82 had done so 2–4 times, and 26 had used them more than four times.

5.3.3 Study Design and Procedure.
We used a between-subjects design with six groups of 30 participants each. Participants were assigned to one of three response variants (RVs) conditions (P1–P30; P31–P60; P61–P90) or one of three feedback variants (FVs) conditions (P91–P120; P121–P150; P151–P180). The response variants did not include any feedback, and the feedback variants did not allow users to re-record or edit their responses. This separation was intentional to avoid interaction effects and to isolate the individual impact of each feature.

Participants were asked to approach the mock interview as if it were a real one, to simulate the sense of sincerity and engagement found in high-stakes settings. To prevent unnecessary stress, we informed them that their responses would not be evaluated. We also provided a briefing that explained what the system would collect, how their data would be used, and what to expect during the study. During the session, participants recorded responses to behavioral interview questions using our system (see Appendix A.4) [80]. Depending on their assigned group, they encountered different interface layouts (Figure 6). After completing the task, they filled out a post-task survey that included Likert scale items (1 = Strongly disagree, 5 = Strongly agree) and an open-ended question about their experience. The entire task was conducted remotely.

5.3.4 Data Analysis.
We conducted Welch ANOVA with Games-Howell post hoc tests to compare Likert-scale responses across interface variants. Two independent coders analyzed the open-ended responses thematically using a grounded theory approach. They began by familiarizing themselves with the data and independently generating initial codes through open coding. They then iteratively refined the codebook through axial coding, consolidating related codes and building a shared understanding of emerging concepts. Using selective coding, they identified the main themes across responses. After establishing the thematic structure, both coders independently applied the agreed-upon themes to

Variable	RV1 (Re-record)	RV2 (Edit)	RV3 (Both Options)	Post-Hoc Analysis
Sense of Control $F(2, 57) = 6.83$, $p = .001$	**M = 3.567**, SD = 0.497	**M = 3.678**, SD = 0.470	M = 3.489, SD = 0.503	RV2 > RV3* RV2 > RV1*
Usefulness $F(2, 57) = 2.62$, $p = .075$	**M = 3.567**, SD = 0.498	M = 3.400, SD = 0.501	M = 3.456, SD = 0.493	–
Confidence $F(2, 56) = 3.65$, $p = .028$	M = 3.322, SD = 0.577	M = 3.500, SD = 0.545	**M = 3.533**, SD = 0.503	RV3 > RV1*
Expressiveness $F(2, 57) = 6.11$, $p = .003$	M = 3.211, SD = 0.743	**M = 3.322**, SD = 0.668	M = 2.967, SD = 0.694	RV2 > RV3** RV1 > RV3*

Table 10. Welch ANOVA and Games–Howell post hoc comparisons across response variants or RVs. Significant post-hoc differences are marked with asterisks (* $p < .05$, ** $p < .01$, *** $p < .001$).

the remaining responses. We assessed inter-rater reliability using Cohen's kappa (RVs: 0.81, FVs: 0.84), which indicated a high level of agreement.

6 Study 2: Result

6.1 Response Interface Variants: Quantitative Analysis

We evaluated the effect of three response variants: RV1 (Re-record), RV2 (Edit), and RV3 (Both) on four outcome variables: sense of control, expressiveness, confidence, and usefulness. A Welch ANOVA was used to test for differences among groups. The results are presented in Table 10.

We found statistically significant differences for sense of control ($F(2, 57) = 6.83$, $p = .001$, $\eta^2 = .047$), confidence ($F(2, 56) = 3.65$, $p = .028$, $\eta^2 = .029$), and expressiveness ($F(2, 57) = 6.11$, $p = .003$, $\eta^2 = .041$). According to the effect sizes, these differences are small to moderate, but meaningful in the context of interaction design.

Post hoc Games–Howell comparisons revealed that for sense of control, participants rated RV2 (Edit) significantly higher than both RV1 (Re-record) and RV3 (Both). For expressiveness, RV2 was again significantly more preferred than RV3, and also showed a modest advantage over RV1. In terms of confidence, RV3 outperformed RV1, but the difference between RV2 and RV3 was not significant.

No significant differences were observed for usefulness ($p = .075$) although descriptively, RV1 scored slightly higher. This suggests that while participants viewed all variants as generally fair and respectful, these dimensions did not vary strongly with interaction design.

6.2 Response Interface Variants: Qualitative Analysis

6.2.1 *Re-record Response (RV1): Simplicity with Repetition Cost.*

Participants using RV1 often commented on its simplicity and predictability. There was no ambiguity in how to proceed: if something went wrong, the only option was to re-record. For some users, this structure was helpful and easy to follow. However, many found it mentally tiring and time-consuming. Small mistakes meant starting over entirely, leading to frustration and disengagement over time. Some participants reported accepting suboptimal recordings rather than starting again, indicating a trade-off between effort and satisfaction. Others felt discouraged by the high stakes attached to minor slip-ups, which reduced their sense of confidence as mentioned by P12, *"Even tiny slip-ups felt costly because there was no way to fix just that part"*. Although the interface was easy to understand, its lack of granularity seemed to amplify emotional pressure. Therefore, RV1 may provide procedural clarity but limits opportunities for low-effort recovery, which can gradually reduce user confidence with repeated use. This aligns with the lowest confidence score for RV1 and its lower sense of control compared to RV2.

6.2.2 *Edit Response (RV2): Clarity, Control, and Unexpected Pressure.*

RV2 was most commonly described as precise, efficient, and satisfying. Participants found editing opportunity helpful for correcting minor errors without needing to repeat their full response. The process felt streamlined and low-effort, especially when compared to re-recording. Many comments reflected that this tool enabled users to make quick adjustments and feel in control of their final message. P33 shared, *"It felt like I had more control over my response, which made the process satisfying."*. This is reflected in RV2's top scores for both sense of control and self-representation. However, a smaller subset of participants used editing not only to correct errors but also to reword or reshape entire thoughts. These users engaged in multiple rounds of revision, not out of necessity, but from a concern about tone, impression, or misinterpretation. For them, editing prompted a type of performative caution, where self-presentation mattered as much as correctness. This behavior introduces a possible tension: while editing increased perceived control, it may also have increased anxiety around how responses might be judged. This may explain why confidence did not differ significantly from RV3, despite RV2's strengths. Thus, although RV2 was most preferred overall, it may also encourage self-censorship or over-editing in evaluative context.

6.2.3 *Both to Choose From (RV3): Freedom with Ambiguity.*

The interface that allowed both editing and re-recording (RV3) received mixed feedback. Some participants appreciated its flexibility and valued having the choice to select the method that best suited their needs. This flexibility may have contributed to RV3's highest confidence rating, significantly outperforming RV1. However, others reported feeling uncertain about which option to use and often second-guessed their decisions. This added an extra layer of friction, particularly in situations where the response felt especially important. Although participants could only use one option per attempt, some described thinking through both before making a decision, which added to their cognitive load. The absence of system guidance or prompts left users to figure out their own approach during the task. In this way, RV3's flexibility came with increased mental effort which is supported by RV3's significantly lower self-representation score compared to RV2, despite offering both options. As P78 put in, *"The flexibility was reassuring, even if I had to think a bit more"*.

6.3 Response Interface Variants: Summary of Quantitative and Qualitative Analysis

Overall, we found that RV2 scored highest on control and expressiveness, and participants attributed this to the ease of fixing small mistakes without repeating entire responses, which made them feel both efficient and well represented. Its confidence score did not exceed RV3 because some users over edited out of concern for how they might appear, adding additional pressure on themselves despite the tool's strengths. RV1 showed the lowest confidence and control, which matches participants' self reports that mandatory re-recording created fatigue and led them to accept imperfect responses simply to avoid extra effort. RV3 achieved the highest confidence because participants liked having a choice, yet its lower expressiveness reflects the uncertainty users described when deciding between editing and re-recording.

6.4 Feedback Interface Variants: Quantitative Analysis

We evaluated the effect of three feedback variants: FV1 (Motivational), FV2 (Evaluative), and FV3 (Both) on five outcome variables: perceived usefulness, confidence, engagement and, acknowledgment. Similar to RVs analysis, Welch ANOVA was used to test for differences across groups. The results are presented in Table 11.

We found statistically significant differences for usefulness ($F(2, 55) = 3.54$, $p = .036$, $\eta^2 = .107$), confidence ($F(2, 51) = 3.79$, $p = .029$, $\eta^2 = .064$), and acknowledgment ($F(2, 57) = 6.30$, $p = .003$, $\eta^2 = .138$). According to the effect sizes, these differences were moderate to large, indicating that

Variable	FV1 (Motivational)	FV2 (Evaluative)	FV3 (Both)	Post-Hoc Analysis
Usefulness $F(2, 55) = 3.54$, $p = .036$	M = 3.467, SD = 0.507	M = 2.967, SD = 1.033	**M = 3.533**, SD = 0.571	FV3 > FV2*
Confidence $F(2, 51) = 3.79$, $p = .029$	**M = 3.300**, SD = 0.535	M = 2.933, SD = 0.980	M = 2.667, SD = 1.348	FV1 > FV3 (marginal significance; $p < .06$)
Engagement $F(2, 58) = 1.46$, $p = .241$	M = 3.333, SD = 0.661	M = 3.300, SD = 0.794	M = 3.033, SD = 0.765	–
Acknowledgment $F(2, 57) = 6.30$, $p = .003$	M = 3.100, SD = 0.662	M = 2.700, SD = 0.837	**M = 3.400**, SD = 0.675	FV3 > FV2**

Table 11. Welch ANOVA and Games–Howell post hoc comparisons across feedback variants or FVs. Significant post-hoc differences are marked with asterisks (* $p < .05$, ** $p < .01$, *** $p < .001$).

the type of feedback meaningfully influenced user perceptions. No significant differences were found for engagement ($p = .241$), suggesting this aspect was less sensitive to variation in feedback design variations.

Post hoc Games–Howell comparisons revealed that for usefulness, participants rated FV3 (Both) significantly higher than FV2 (Evaluative) ($p = .031$), while FV1 (Motivational) also showed a marginal advantage over FV2 ($p = .056$). For confidence, FV1 was marginally preferred over FV3 ($p = .055$), indicating that motivational feedback alone may be perceived as more effective than when combined with evaluative cues to enhance competence. For acknowledgment, FV3 was rated significantly higher than FV2 ($p = .002$), highlighting the benefit of combining motivational and evaluative elements for offering a sense of being valued.

No significant pairwise differences were found for engagement, and the mean values were relatively similar across all feedback variants. This suggests that the type of feedback provided may not have a strong influence on enhancing participants' sense of engagement to perform better.

6.5 Feedback Interface Variants: Qualitative Analysis

6.5.1 Motivational Feedback (FV1): Emotional Safety Interpreted as Presence.

User responses to the motivational feedback condition indicated a consistent perception of emotional support and attentiveness. Although this variant offered only affirmation and lacked performance-related critique, participants rated it highly on perceived confidence. This pattern suggests that motivational feedback was not interpreted as superficial or generic. Rather, the absence of criticism was perceived as intentional and supportive. Beyond affective reception, it may also have encouraged greater self-regulation. By not prescribing specific changes, the system implicitly placed the responsibility on the user to reflect and adjust on their own terms. In this way, the design may have supported autonomy by offering space for self-directed improvement rather than externally guided correction. As P107 described *"It (feedback) felt like someone cheering me on while I was answering [...]"* and P116 found it *"warm and supportive"* reinforcing the sense of emotional alignment.

6.5.2 Evaluative Feedback (FV2): Constructive Yet Emotionally Distant.

This variant delivered targeted, human-like suggestions after each response and was clearly personalized at the task level. However, it was rated lowest in acknowledgment and usefulness. While the feedback was informative, its tone appeared to prioritize correction over connection. Participants (P124, P139 & P147) described it as *"critical"*, indicating that the system's emphasis on identifying areas for improvement may have diminished the sense of relational presence. P141 mentioned *"It gave useful suggestions but [...] felt more like an evaluation than a conversation"*. Despite being constructive, the feedback may have been perceived as judgmental rather than supportive. The experience of repeated critique, even when phrased with care, appeared to create a sense of being

monitored rather than guided. In offering direction, the design may have compromised users' sense of recognition.

6.5.3 Both Feedback Types (FV3): Acknowledging Without Anchoring.
The combined motivational and evaluative feedback after each response received the highest ratings in acknowledgment and usefulness. Although it suggests that users felt both recognized and guided, its perceived confidence score was lower than motivational feedback, indicating some ambiguity in how the feedback was interpreted. While the balance of encouragement and critique may have helped reduce emotional strain, it may also have made the system's intent less clear. Notably, this condition received the lowest mean score in engagement. Although the difference was not statistically significant, the trend suggests that incorporating evaluative elements may have diluted the emotional impact that motivational feedback alone achieved, as reported by P159: *"supportive, but also pointed out too many things"*. At the same time, the strong ratings in acknowledgment and usefulness indicate that combining critique with support may have addressed a broader range of user expectations as P179 commented, *"It gave me direction without feeling harsh"*. Also, participants (P162 & P171) perceived it as a form of *"gentle criticism"* and beneficial for their self-improvement.

6.6 Feedback Interface Variants: Summary of Quantitative and Qualitative Analysis

We found that FV3 scored highest on usefulness and acknowledgment, and participants attributed this to receiving both encouragement and concrete guidance, which made them feel recognized and supported while still knowing how to improve. Its confidence score did not exceed FV1 because some users felt that mixing critique with praise created uncertainty about the system's intent, adding mild interpretive pressure despite its strong benefits. FV2 showed the lowest usefulness and acknowledgment, which aligns with participants comments that predominantly evaluative messages felt critical and emotionally distant, leading them to feel corrected rather than valued. FV1 achieved the highest confidence because participants interpreted motivational feedback as warm and attentive, yet its lower usefulness reflects the limited direction users described when the system offered affirmation without specific suggestions.

7 Discussion

7.1 Study 1: Performing for the Machine at the Cost of Authenticity

Participants approached AI interviews with expectations shaped by exposure to current AI-mediated technologies and by inflated organizational terminology. Familiarity with tools like ChatGPT facilitated expectations of responsiveness and intelligence, while corporate terms such as "AI recruiter" implied advanced, conversational systems. However, the reality differed: static interfaces, no feedback, and rigid processes that felt indifferent rather than adaptive. This gap reflects Expectation Violation Theory (EVT), in which anticipated social dynamics fail to occur, producing frustration, detachment, or mistrust [22, 54]. Such misalignment often leads to negative reactions, especially when systems fail to meet expectations set by their design or framing [83, 120].

The disappointment was not only about functionality but also about recognition. Participants carefully managed their impressions to match imagined algorithmic preferences. Studies show that algorithmic assessments frequently limit self-expression and ignore individual context [4]. Many began viewing the process as theatrical, even dehumanizing; "not even real," one Redditor said. Others questioned whether the interview was a legitimate evaluation or merely a pipeline for training data. This skepticism echoes broader findings that candidates often feel emotionally misjudged or unseen by automated systems, especially in asynchronous formats [84, 115]. The lack of transparency and organizational silence, including ghosting, deepened this suspicion. Ghosting is not uncommon in hiring, but in this context, where no human presence was ever felt, it did not just

signal rejection; it made participants feel invisible, as if their effort had never mattered at all. Such dehumanizing perceptions are well documented in responses to emotion AI and hiring automation, where the absence of human acknowledgment leads to moral disengagement and reduced trust [109, 120]. In response, applicants described using AI tools to script answers, altering speech delivery, or uploading polished audio. This behavior reflects strategic adaptation to the interviewer algorithm, aligning with prior research on workarounds and appropriation in sociotechnical systems [33, 40].

For neurodivergent participants, the stakes were higher. Research shows that marginalized users often hide identity traits to avoid misinterpretation by algorithmic systems, which contributes to systemic inequity [134, 144]. If applicants feel compelled to mask their neurodivergence or find workarounds to appear "neutral," the system no longer supports equity; it instead rewards conformity. The case of AI interviews adds another dimension: systems designed for efficiency and standardization can inadvertently enforce a "one-size-fits-all" interaction style, which can be challenging for neurodivergent users and others whose communication practices diverge from the norm. Prior work underscores the need to design AI systems that are inclusive of neurodivergent users, not just accessible in theory but also responsive in practice [144].

7.2 Study 2: Designing System Toward Equitable Experiences

Our findings highlight how user agency and transparency shape the candidate experience in AI-mediated interviews, particularly through response control and feedback. Response variants provide candidates with some control over self-presentation, but this control is uneven and often difficult to interpret. Re-recording allows candidates to start over; however, it can feel demanding because candidates must improve their responses without interactional guidance. Prior work shows that many challenges in asynchronous AI interviews stem from the absence of interactional cues that signal understanding and evaluative stance in human interviews [119]. Without in-the-moment feedback, candidates lack information about how their responses are received, which limits their ability to adjust or clarify and reduces perceived control.

Editing within a limited window offers a more bounded form of agency, allowing candidates to refine responses without discarding them entirely. However, flexibility alone does not ensure a sense of control. When options are not clearly framed, flexibility can increase uncertainty [4], as candidates question which choice aligns with system expectations. Prior work suggests that candidates often seek clearer opportunities to explain their intent or provide context, such as clarifying circumstances or technical issues, rather than having unlimited control [115]. This points to design directions that emphasize transparent and constrained forms of agency, where candidates understand what can be changed and how those changes are used.

Feedback further shapes perceptions of control and openness. Motivational feedback was often experienced as supportive, while evaluative feedback provided direction but could feel intrusive if delivered during the interaction. Perhaps feedback is better received after the interview, where it supports reflection rather than immediate judgment [107, 125]. From a transparency perspective, feedback signals how the system and the organization interpret candidate responses. Generic feedback may offer reassurance, but more contextual feedback can help candidates better understand evaluation standards and outcomes [6].

Recruiters and HR stakeholders further emphasize that organizational practices play a significant role in shaping how agency and transparency are experienced in AI-mediated interviews. While these systems are often valued for efficiency, stakeholders acknowledge that candidates have limited visibility into how responses are assessed, particularly in asynchronous settings [87, 140]. As a result, candidates often rely on inference rather than explicit communication to understand expectations or next steps in the process [66]. Stakeholders note that decisions about response

flexibility and feedback are typically determined through employer-controlled system settings, such as whether revisions are allowed and how many attempts are permitted. Although these configurations are intended to balance consistency and fairness, the reasoning behind them and their role in evaluation are rarely communicated to candidates, which can make it difficult for candidates to understand the boundaries of their control.

At the same time, employers commonly frame AI interview systems as tools that support, rather than replace, human decision-making [117]. From the candidate perspective, however, this distinction is not always apparent during the interview itself. Prior organizational research suggests that how AI systems are embedded in practice influences how judgments are formed and how accountability is perceived, indicating that transparency depends not only on system design but also on how organizations communicate their use of these tools [62, 108]. Taken together, these perspectives suggest that user agency and transparency in AI-mediated interviews are shaped through both interface features and organizational communication. Stakeholders therefore point to the need for approaches that make the boundaries of system control more visible and that more clearly communicate expectations, interpretation, and follow-up, rather than relying on technical improvements alone [43, 88].

7.3 From Inflated Expectations to Algorithmic Sensemaking

The growing advancement of technologies has raised the bar for user expectations. As interactions with advanced systems become routine, users internalize certain standards, making it difficult to accept systems that fall short. With increasing reliance on and extensive use of LLM-driven systems [60, 70, 73], users are beginning to expect highly conversational and personalized interactions. This marks a shift in anthropomorphism, where users increasingly compare automated systems not to humans but to advanced language models. As others have observed, users may place undue trust in LLMs when they appear more human, influencing perceptions of credibility and authority [30]. Expectations and interpretations often arise from one-sided experiences, where users approach systems with preformed assumptions that become more pronounced in the absence of transparent design. The interpretation of black-box systems is often shaped by subjective experience when objective explanations are limited or unavailable [9, 45]. While such subjectivity may appear trivial, our findings show that it can lead to serious trust breakdowns, such as when system opacity and rigidity lead applicants to suspect that the platform is a data-harvesting scam rather than a legitimate interview system.

Beyond conversational use, AI-driven systems support a wide range of functions across domains. When organizations report using AI-driven systems, users often associate them primarily with interaction, even though such systems may be deployed solely for analysis or decision support [122, 145]. Adopting AI-based technologies also remains costly for many small and mid-sized enterprises [74]. While these constraints are understandable, organizations are still expected to demonstrate care toward applicants, which is often absent in practice despite promises to the contrary. Prior research shows that hiring services frequently make unsupported claims about solving hiring-related problems through their technologies, many of which do not hold up under scrutiny [120], and that such tools are often deployed without thorough evaluation [128]. In this context, our findings show that candidates adapt to uncertainty by forming folk theories about what the system values and by adjusting their tone and wording to perform for an unseen algorithm. This aligns with prior work on algorithmic sensemaking [36, 151] and with research showing that the structure of an evaluative environment can shape participants' behavior [139]. Over time, this adaptation can narrow acceptable forms of self-presentation, encouraging safe and generic responses while discouraging personal or creative expression, echoing concerns that

opaque systems promote risk-averse behavior [17]. These dynamics heighten the risk of distrust and unmet expectations driven by inflated claims [2], and in some cases may also perpetuate bias.

The findings further point to the limits of interface-level interventions in shaping candidate experience. While edit options reduced cognitive and emotional burden by lowering the perceived risk of being candid, performance-oriented strategies largely persisted, suggesting that interaction design alone may not fully address deeper concerns about evaluation and recognition. Prior ACM research shows that perceptions of fairness and legitimacy in algorithmic decision-making depend less on transparency alone and more on signals of oversight, contestability, and accountability [91]. At the same time, organizational constraints and efficiency pressures limit how much agency AI-driven hiring systems can realistically offer applicants. Given that applicants occupy a less powerful position and often operate under constrained conditions, responsibility falls more heavily on organizations to ensure that hiring processes enable fair and authentic performance [89]. More broadly, AI-driven interview tools are expected to support fairness, consistency, and accessibility. CSCW research has long emphasized empowering users to recognize and resist manipulative patterns and designing systems that uphold justice by treating individuals fairly across identities and contexts [1, 100, 115, 136]. Embedding these values into system design is therefore not merely a theoretical concern but a practical requirement for building sustainable and equitable hiring technologies.

8 Limitations and Future Work

While our work offers important contributions to the study of AI-driven hiring, it is not without limitations. First, although we analyzed a large volume of *Reddit* data, the authenticity of these posts could not be independently verified. Similarly, our user studies relied on self-reported data, which may be subject to bias or inaccuracies. Second, the majority of our participants were students, which limits the generalizability of our findings to broader populations. Third, Study 2 involved a simulated interview scenario, which may not fully reflect real-world stakes or pressures. The transparency provided during the study may also have influenced participants' responses in ways that differ from commercial deployments. Fourth, our work centers primarily on applicant perspectives, omitting the organizational viewpoint. Understanding how employers implement, interpret, and evaluate these systems remains an important area for future work. Fifth, while we identified several tools marketed as AI-driven based on company websites, we could not verify whether the organizations using them employed the same versions or configurations.

In future work, we plan to recruit professionals with industry experience, hiring managers and organizations to deepen our understanding of AI interview practices in operational contexts. We also intend to include participants from more diverse age groups and occupations to increase ecological validity.

9 Conclusion

This paper advances the understanding of AI-driven asynchronous interviews by examining how job seekers experience and adapt to these systems amid the growing public familiarity with LLMs. Through large-scale analysis of *Reddit* discussions and follow-up interviews, we uncovered critical mismatches between system design and user expectations mismatches often rooted in organizational rhetoric, anthropomorphic interface cues, and evolving assumptions about AI capabilities. These mismatches manifested as diminished trust, reduced perceived autonomy, and a sense of disconnection, among both neurodivergent and neurotypical users. Building on these findings, we designed an interview interface and evaluated based on Self-Determination Theory (SDT). By offering different configurations of response control and feedback, we explored how design can influence users' sense of agency, competence, and acknowledgment. Our results demonstrate that

simple yet intentional features such as editing options and motivational feedback can significantly shape candidate experience, while overly complex or ambiguous combinations may undermine user confidence. Together, these findings underscore the importance of expectation management, sense of agency, and psychological support in the design of AI hiring systems.

References

[1] Yalcin Acikgoz, Kristl H Davison, Maira Compagnone, and Matt Laske. 2020. Justice perceptions of artificial intelligence in selection. *International Journal of Selection and Assessment* 28, 4 (2020), 399–416.

[2] Rudaiba Adnin and Maitraye Das. 2024. " I look at it as the king of knowledge": How Blind People Use and Understand Generative AI Tools. In *Proceedings of the 26th International ACM SIGACCESS Conference on Computers and Accessibility*. 1–14.

[3] MIT: Career Advising and Professional Development. 2020. Using the STAR method for your next behavioral interview. https://capd.mit.edu/resources/the-star-method-for-behavioral-interviews/

[4] Evgeni Aizenberg, Matthew J Dennis, and Jeroen van den Hoven. 2025. Examining the assumptions of AI hiring assessments and their impact on job seekers' autonomy over self-representation. *AI & society* 40, 2 (2025), 919–927.

[5] Alicia C Alonzo and Jiwon Kim. 2016. Declarative and dynamic pedagogical content knowledge as elicited through two video-based interview methods. *Journal of Research in Science Teaching* 53, 8 (2016), 1259–1286.

[6] Mike Ananny and Kate Crawford. 2018. Seeing without knowing: Limitations of the transparency ideal and its application to algorithmic accountability. *new media & society* 20, 3 (2018), 973–989.

[7] Nazanin Andalibi, Luke Stark, Daniel McDuff, Rosalind Picard, Jonathan Gratch, and Noura Howell. 2024. What should we do with Emotion AI? Towards an Agenda for the Next 30 Years. In *Companion Publication of the 2024 Conference on Computer-Supported Cooperative Work and Social Computing*. 98–101.

[8] Zinat Ara, Amrita Ganguly, Donna Peppard, Dongjun Chung, Slobodan Vucetic, Vivian Genaro Motti, and Sungsoo Ray Hong. 2024. Collaborative job seeking for people with autism: Challenges and design opportunities. In *Proceedings of the 2024 CHI conference on human factors in computing systems*. 1–17.

[9] Lena Armstrong, Jayne Everson, and Amy J Ko. 2023. Navigating a black box: Students' experiences and perceptions of automated hiring. In *Proceedings of the 2023 ACM Conference on International Computing Education Research-Volume 1*. 148–158.

[10] Lena Armstrong and Danaé Metaxa. 2025. Navigating Automated Hiring: Perceptions, Strategy Use, and Outcomes Among Young Job Seekers. *Proceedings of the ACM on Human-Computer Interaction* 9, 2 (2025), 1–26.

[11] Nagadivya Balasubramaniam, Marjo Kauppinen, Antti Rannisto, Kari Hiekkanen, and Sari Kujala. 2023. Transparency and explainability of AI systems: From ethical guidelines to requirements. *Information and Software Technology* 159 (2023), 107197.

[12] Nick Ballou, Sebastian Deterding, April Tyack, Elisa D Mekler, Rafael A Calvo, Dorian Peters, Gabriela Villalobos-Zúñiga, and Selen Turkay. 2022. Self-determination theory in HCI: shaping a research agenda. In *CHI conference on human factors in computing systems extended abstracts*. 1–6.

[13] BBC. 2022. Job hunting for neurodivergent people: 'AI recruitment means I've got zero chance'. https://www.bbc.com/bbcthree/article/c62bcab6-db6f-4026-90bb-7f508705a65b

[14] Daniel Bennett and Elisa D Mekler. 2024. Beyond Intrinsic Motivation: The Role of Autonomous Motivation in User Experience. *ACM Transactions on Computer-Human Interaction* 31, 5 (2024), 1–41.

[15] Nanyi Bi and Janet Yi-Ching Huang. 2023. I create, therefore I agree: Exploring the effect of AI anthropomorphism on human decision-making. In *Companion Publication of the 2023 Conference on Computer Supported Cooperative Work and Social Computing*. 241–244.

[16] Austin Biehl, Daniela Perez, Jessica Ingle, and Morgan G Ames. 2025. Prestige and Prejudice: How the Interplay of Recruiting Work and Algorithms Reinforces Social Inequities in Software Engineering. In *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems*. 1–16.

[17] Reuben Binns, Max Van Kleek, Michael Veale, Ulrik Lyngs, Jun Zhao, and Nigel Shadbolt. 2018. 'It's Reducing a Human Being to a Percentage' Perceptions of Justice in Algorithmic Decisions. In *Proceedings of the 2018 Chi conference on human factors in computing systems*. 1–14.

[18] Shreyan Biswas, Ji-Youn Jung, Abhishek Unnam, Kuldeep Yadav, Shreyansh Gupta, and Ujwal Gadiraju. 2024. "Hi. I'm Molly, Your Virtual Interviewer!" Exploring the Impact of Race and Gender in AI-Powered Virtual Interview Experiences. In *Proceedings of the AAAI Conference on Human Computation and Crowdsourcing*, Vol. 12. 12–22.

[19] Karen L Boyd and Nazanin Andalibi. 2023. Automated emotion recognition in the workplace: How proposed technologies reveal potential futures of work. *Proceedings of the ACM on human-computer interaction* 7, CSCW1 (2023), 1–37.

[20] Virginia Braun and Victoria Clarke. 2006. Using thematic analysis in psychology. *Qualitative research in psychology* 3, 2 (2006), 77–101.
[21] Virginia Braun and Victoria Clarke. 2019. Reflecting on reflexive thematic analysis. *Qualitative research in sport, exercise and health* 11, 4 (2019), 589–597.
[22] Judee K Burgoon. 2015. Expectancy violations theory. *The international encyclopedia of interpersonal communication* (2015), 1–9.
[23] Maarten Buyl, Christina Cociancig, Cristina Frattone, and Nele Roekens. 2022. Tackling algorithmic disability discrimination in the hiring process: An ethical, legal and technical analysis. In *Proceedings of the 2022 ACM Conference on Fairness, Accountability, and Transparency*. 1071–1082.
[24] Peter Cappelli. 2019. Your approach to hiring is all wrong. *Harvard Business Review* 97, 3 (2019), 48–58.
[25] Nancy Carter. 2014. The use of triangulation in qualitative research. *Number 5/September 2014* 41, 5 (2014), 545–547.
[26] Ruijia Cheng, Ziwen Zeng, Maysnow Liu, and Steven Dow. 2020. Critique me: exploring how creators publicly request feedback in an online critique community. *Proceedings of the ACM on Human-Computer Interaction* 4, CSCW2 (2020), 1–24.
[27] Phoebe K Chua and Melissa Mazmanian. 2020. Are you one of us? Current hiring practices suggest the potential for class biases in large tech companies. *Proceedings of the ACM on Human-Computer Interaction* 4, CSCW2 (2020), 1–20.
[28] Phoebe K Chua and Melissa Mazmanian. 2022. The substance of style: How social class-based styles of interpersonal interaction shape hiring assessments at large technology companies. *Proceedings of the ACM on Human-Computer Interaction* 6, CSCW2 (2022), 1–22.
[29] Louis Cohen, Lawrence Manion, and Keith Morrison. 2017. Coding and content analysis. In *Research methods in education*. Routledge, 668–685.
[30] Michelle Cohn, Mahima Pushkarna, Gbolahan O Olanubi, Joseph M Moran, Daniel Padgett, Zion Mengesha, and Courtney Heldreth. 2024. Believing anthropomorphism: examining the role of anthropomorphic cues on trust in large language models. In *Extended Abstracts of the CHI Conference on Human Factors in Computing Systems*. 1–15.
[31] Shanley Corvite, Kat Roemmich, Tillie Ilana Rosenberg, and Nazanin Andalibi. 2023. Data subjects' perspectives on emotion artificial intelligence use in the workplace: A relational ethics lens. *Proceedings of the ACM on Human-Computer Interaction* 7, CSCW1 (2023), 1–38.
[32] Carolyn Crist. 2024. CVS settles lawsuit alleging it used AI 'lie detector' in interviews. https://www.hrdive.com/news/cvs-settles-lawsuit-over-using-ai-based-lie-detector/722249/
[33] Robert M Davison, Louie HM Wong, Steven Alter, and Carol XJ Ou. 2019. Adopted globally but unusable locally: What workarounds reveal about adoption, resistance, compliance and noncompliance. In *27th European Conference on Information Systems (ECIS 2019): Information Systems for a Sharing Society*. Association for Information Systems.
[34] Stefano De Paoli. 2023. Performing an Inductive Thematic Analysis of Semi-Structured Interviews With a Large Language Model: An Exploration and Provocation on the Limits of the Approach. *Social Science Computer Review* (2023), 08944393231220483.
[35] Alaina Demopoulos. 2024. The Gurdian. Retrieved March 06, 2020 from https://www.theguardian.com/technology/2024/mar/06/ai-interviews-job-applications
[36] Michael Ann DeVito. 2021. Adaptive folk theorization as a path to algorithmic literacy on changing platforms. *Proceedings of the ACM on Human-Computer Interaction* 5, CSCW2 (2021), 1–38.
[37] Hitesh Dhiman, Gustavo Alberto Rovelo Ruiz, Raf Ramakers, Danny Leen, and Carsten Röcker. 2024. Designing instructions using self-determination theory to improve motivation and engagement for learning craft. In *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems*. 1–16.
[38] Tawanna R Dillahunt and Joey Chiao-Yin Hsiao. 2020. Positive feedback and self-reflection: features to support self-efficacy among underrepresented job seekers. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. 1–13.
[39] Robert L Dipboye, Therese Macan, and Comila Shahani-Denning. 2012. The selection interview from the interviewer and applicant perspectives: Can't have one without the other. *The Oxford handbook of personnel assessment and selection* (2012), 323–352.
[40] Olivia Doggett, Sarah Gram, Sophia Jit, Basel Harb, Houssam Harb, Mohamad Shafik, Robert Soden, and Matt Ratto. 2024. Digital Gig Work Under Unstable Energy Infrastructures: Invisible Workarounds and Alternative Imaginaries. *Proceedings of the ACM on Human-Computer Interaction* 8, CSCW2 (2024), 1–26.
[41] Liz Dowthwaite, Alexa Spence, Chris Lintott, Grant Miller, James Sprinks, Sam Blickhan, and Robert Houghton. 2025. Exploring the Relationship between Basic Psychological Needs and Motivation in Online Citizen Science. *ACM Transactions on Social Computing* 8, 1-2 (2025), 1–36.
[42] Alessandro Fabris, Nina Baranowska, Matthew J Dennis, David Graus, Philipp Hacker, Jorge Saldivar, Frederik Zuiderveen Borgesius, and Asia J Biega. 2025. Fairness and bias in algorithmic hiring: A multidisciplinary survey. *ACM Transactions on Intelligent Systems and Technology* 16, 1 (2025), 1–54.

[43] Daniel Felz and Kimberly Peretti. 2022. AI Regulation in the US: What's Coming and What Companies Need to Do in 2023. Newsletter.
[44] Uwe Flick. 2004. Triangulation in qualitative research. *A companion to qualitative research* 3 (2004), 178–183.
[45] Ruth C Fong and Andrea Vedaldi. 2017. Interpretable explanations of black boxes by meaningful perturbation. In *Proceedings of the IEEE international conference on computer vision*. 3429–3437.
[46] Eureka Foong, Steven P Dow, Brian P Bailey, and Elizabeth M Gerber. 2017. Online feedback exchange: A framework for understanding the socio-psychological factors. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*. 4454–4467.
[47] Denae Ford, Titus Barik, Leslie Rand-Pickett, and Chris Parnin. 2017. The tech-talk balance: what technical interviewers expect from technical candidates. In *2017 IEEE/ACM 10th International Workshop on Cooperative and Human Aspects of Software Engineering (CHASE)*. IEEE, 43–48.
[48] Chengguang Gan, Qinghao Zhang, and Tatsunori Mori. 2024. Application of llm agents in recruitment: A novel framework for resume screening. *arXiv preprint arXiv:2401.08315* (2024).
[49] Radhika Garg, Yash Kapadia, and Subhasree Sengupta. 2021. Using the lenses of emotion and support to understand unemployment discourse on Reddit. *Proceedings of the ACM on Human-Computer Interaction* 5, CSCW1 (2021), 1–24.
[50] Sahin Cem Geyik, Stuart Ambler, and Krishnaram Kenthapadi. 2019. Fairness-aware ranking in search & recommendation systems with application to linkedin talent search. In *Proceedings of the 25th acm sigkdd international conference on knowledge discovery & data mining*. 2221–2231.
[51] Manuel F Gonzalez, Weiwei Liu, Lei Shirase, David L Tomczak, Carmen E Lobbe, Richard Justenhoven, and Nicholas R Martin. 2022. Allying with AI? Reactions toward human-based, AI/ML-based, and augmented hiring processes. *Computers in Human Behavior* 130 (2022), 107179.
[52] Priyanshul Govil, Hemang Jain, Vamshi Bonagiri, Aman Chadha, Ponnurangam Kumaraguru, Manas Gaur, and Sanorita Dey. 2025. Cobias: Assessing the contextual reliability of bias benchmarks for language models. In *Proceedings of the 17th ACM Web Science Conference 2025*. 460–471.
[53] Etienne Grenier and Nicolas Chartier-Edwards. 2024. Automated hiring platforms as mythmaking machines and their symbolic economy. *Journal of Digital Social Research* 6, 4 (2024), 34–48.
[54] G Mark Grimes, Ryan M Schuetzler, and Justin Scott Giboney. 2021. Mental models and expectation violations in conversational AI interactions. *Decision Support Systems* 144 (2021), 113515.
[55] Meghna Gupta, Arpita Bhattacharya, and Julie Kientz. 2025. " Being a nanny isn't just caregiving": An Analysis of How Nannies Seek Support in Online Communities like/r/Nanny. In *Proceedings of the 4th Annual Symposium on Human-Computer Interaction for Work*. 1–15.
[56] Pooja Gupta, Semila F Fernandes, and Manish Jain. 2018. Automation in recruitment: a new frontier. *Journal of Information Technology Teaching Cases* 8, 2 (2018), 118–125.
[57] Hyerim Han, Bogyeom Park, and Kyoungwon Seo. 2025. A Self-Determination Theory-based Career Counseling Chatbot: Motivational Interactions to Address Career Decision-Making Difficulties and Enhance Engagement. In *Proceedings of the Extended Abstracts of the CHI Conference on Human Factors in Computing Systems*. 1–9.
[58] Béatrice S Hasler, Peleg Tuchman, and Doron Friedman. 2013. Virtual research assistants: Replacing human interviewers by automated avatars in virtual worlds. *Computers in Human Behavior* 29, 4 (2013), 1608–1616.
[59] Lobna Hassan, Antonio Dias, and Juho Hamari. 2019. How motivational feedback increases user's benefits and continued use: A study on gamification, quantified-self and social networking. *International Journal of Information Management* 46 (2019), 151–162.
[60] Jessica He, Stephanie Houde, Gabriel E Gonzalez, Darío Andrés Silva Moran, Steven I Ross, Michael Muller, and Justin D Weisz. 2024. AI and the Future of Collaborative Work: Group Ideation with an LLM in a Virtual Canvas. In *Proceedings of the 3rd Annual Meeting of the Symposium on Human-Computer Interaction for Work*. 1–14.
[61] Catherine M Hicks, Vineet Pandey, C Ailie Fraser, and Scott Klemmer. 2016. Framing feedback: Choosing review environment features that support high quality peer assessment. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*. 458–469.
[62] Tad Hirsch, Kritzia Merced, Shrikanth Narayanan, Zac E Imel, and David C Atkins. 2017. Designing contestability: Interaction design, machine learning, and mental health. In *Proceedings of the 2017 Conference on Designing Interactive Systems*. 95–99.
[63] Piotr Horodyski. 2023. Recruiter's perception of artificial intelligence (AI)-based tools in recruitment. *Computers in Human Behavior Reports* 10 (2023), 100298.
[64] Anna Lena Hunkenschroer and Alexander Kriebitz. 2023. Is AI recruiting (un) ethical? A human rights perspective on the use of AI for hiring. *AI and Ethics* 3, 1 (2023), 199–213.
[65] Angel Hsing-Chi Hwang and Andrea Stevenson Won. 2022. Ai in your mind: Counterbalancing perceived agency and experience in human-ai interaction. In *Chi conference on human factors in computing systems extended abstracts*. 1–10.

[66] Ümit Deniz İlhan, Burcu Kümbül Güler, Dilara Turgut, and Cem Duran. 2025. Opportunities and challenges of asynchronous video interviews: Perceptions of human resources professionals from Türkiye. *PLoS One* 20, 6 (2025), e0325932.
[67] Sana Hassan Imam, Christopher A Metz, Lars Hornuf, and Rolf Drechsler. 2024. Determining the Effect of Feedback Quality on User Engagement on Online Idea Crowdsourcing Platforms Using an AI Model. *Proceedings of the ACM on Human-Computer Interaction* 8, CSCW2 (2024), 1–26.
[68] Jonas Ivarsson and Oskar Lindwall. 2023. Suspicious minds: the problem of trust and conversational agents. *Computer Supported Cooperative Work (CSCW)* 32, 3 (2023), 545–571.
[69] Maia Jacobs, Jeffrey He, Melanie F. Pradier, Barbara Lam, Andrew C Ahn, Thomas H McCoy, Roy H Perlis, Finale Doshi-Velez, and Krzysztof Z Gajos. 2021. Designing AI for trust and collaboration in time-constrained medical decisions: a sociotechnical lens. In *Proceedings of the 2021 chi conference on human factors in computing systems*. 1–14.
[70] Maurice Jakesch, Advait Bhat, Daniel Buschek, Lior Zalmanson, and Mor Naaman. 2023. Co-writing with opinionated language models affects users' views. In *Proceedings of the 2023 CHI conference on human factors in computing systems*. 1–15.
[71] Yuan Jia, Bin Xu, Yamini Karanam, and Stephen Voida. 2016. Personality-targeted gamification: a survey study on personality traits and motivational affordances. In *Proceedings of the 2016 CHI conference on human factors in computing systems*. 2001–2013.
[72] Ruth Ellen Jones and Kareem R Abdelfattah. 2020. Virtual interviews in the era of COVID-19: a primer for applicants. *Journal of surgical education* 77, 4 (2020), 733–734.
[73] Shivani Kapania, Ruiyi Wang, Toby Jia-Jun Li, Tianshi Li, and Hong Shen. 2025. 'I'm Categorizing LLM as a Productivity Tool': Examining Ethics of LLM Use in HCI Research Practices. *Proceedings of the ACM on Human-Computer Interaction* 9, 2 (2025), 1–26.
[74] Pavel Krumkachev Karthik Ramachandran, Varun Dhir. 2021. Cloud helps accelerate midsize companies' AI adoption. https://www2.deloitte.com/us/en/insights/industry/technology/ai-adoption-mid-size-companies.html
[75] Judy Kendall. 1999. Axial coding and the grounded theory controversy. *Western journal of nursing research* 21, 6 (1999), 743–757.
[76] Shahedul Huq Khandkar. 2009. Open coding. *University of Calgary* 23, 2009 (2009), 2009.
[77] Da-jung Kim and Youn-kyung Lim. 2019. Co-performing agent: Design for building user-agent partnership in learning and adaptive services. In *Proceedings of the 2019 CHI conference on human factors in computing systems*. 1–14.
[78] John Kim. 2024. Best Reddit communities for recruiters. https://www.paraform.com/blog/best-reddit-communities-for-recruiters
[79] Jin-Young Kim and WanGyu Heo. 2022. Artificial intelligence video interviewing for employment: perspectives from applicants, companies, developer and academicians. *Information Technology & People* 35, 3 (2022), 861–878.
[80] Lewis Kiptanui. 2025. 10 Behavioral Interview Questions (With Sample Answers). https://www.indeed.com/career-advice/interviewing/behavioral-interview-questions
[81] Param Kothari, Paras Mehta, Srushti Patil, and Varsha Hole. 2024. InterviewEase: AI-powered interview assistance. (2024).
[82] J Richard Landis and Gary G Koch. 1977. The measurement of observer agreement for categorical data. *biometrics* (1977), 159–174.
[83] Markus Langer, Tim Hunsicker, Tina Feldkamp, Cornelius J König, and Nina Grgić-Hlača. 2022. "Look! It'sa computer program! It's an algorithm! It's AI!": Does terminology affect human perceptions and evaluations of algorithmic decision-making systems?. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*. 1–28.
[84] Markus Langer, Cornelius J König, and Victoria Hemsing. 2020. Is anybody listening? The impact of automatically evaluated job interviews on impression management and applicant reactions. *Journal of Managerial Psychology* 35, 4 (2020), 271–284.
[85] Markus Langer, Cornelius J König, and Maria Papathanasiou. 2019. Highly automated job interviews: Acceptance under the influence of stakes. *International Journal of Selection and Assessment* 27, 3 (2019), 217–234.
[86] Markus Langer, Cornelius J König, Diana Ruth-Pelipez Sanchez, and Sören Samadi. 2020. Highly automated interviews: Applicant reactions and the organizational context. *Journal of Managerial Psychology* 35, 4 (2020), 301–314.
[87] Mitra Lashkari and Jinghui Cheng. 2023. "Finding the Magic Sauce": Exploring Perspectives of Recruiters and Job Seekers on Recruitment Bias and Automated Tools. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*. 1–16.
[88] Siddique Latif, Hafiz Shehbaz Ali, Muhammad Usama, Rajib Rana, Björn Schuller, and Junaid Qadir. 2022. Ai-based emotion recognition: Promise, peril, and prescriptions for prosocial path. *arXiv preprint arXiv:2211.07290* (2022).
[89] Maude Lavanchy, Patrick Reichert, Jayanth Narayanan, and Krishna Savani. 2023. Applicants' fairness perceptions of algorithm-driven hiring procedures. *Journal of Business Ethics* 188, 1 (2023), 125–150.

[90] BC Lee and BY Kim. 2021. Development of an AI-based interview system for remote hiring. *International Journal of Advanced Research in Engineering and Technology (IJARET)* 12, 3 (2021), 654–663.
[91] Min Kyung Lee, Daniel Kusbit, Evan Metsky, and Laura Dabbish. 2015. Working with machines: The impact of algorithmic and data-driven management on human workers. In *Proceedings of the 33rd annual ACM conference on human factors in computing systems*. 1603–1612.
[92] Lan Li, Tina Lassiter, Joohee Oh, and Min Kyung Lee. 2021. Algorithmic hiring in practice: Recruiter and HR Professional's perspectives on AI use in hiring. In *Proceedings of the 2021 AAAI/ACM Conference on AI, Ethics, and Society*. 166–176.
[93] Yuan Li, Lichao Sun, and Yixuan Zhang. 2025. MetaAgents: Large Language Model Based Agents for Decision-making on Teaming. *Proceedings of the ACM on Human-Computer Interaction* 9, 2 (2025), 1–27.
[94] Bingjie Liu, Lewen Wei, Mu Wu, and Tianyi Luo. 2023. Speech production under uncertainty: how do job applicants experience and communicate with an AI interviewer? *Journal of Computer-Mediated Communication* 28, 4 (2023), zmad028.
[95] Zhe Liu. 2025. Interview AI-ssistant: Designing for Real-Time Human-AI Collaboration in Interview Preparation and Execution. In *Companion Proceedings of the 30th International Conference on Intelligent User Interfaces*. 225–228.
[96] Eden-Raye Lukacik, Joshua S Bourdage, and Nicolas Roulin. 2022. Into the void: A conceptual model and research agenda for the design and use of asynchronous video interviews. *Human Resource Management Review* 32, 1 (2022), 100789.
[97] Brian D Lyons, Robert H Moorman, and John W Michel. 2024. The vanishing applicant: Uncovering aberrant antecedents to ghosting behaviour. *Journal of Occupational and Organizational Psychology* 97, 4 (2024), 1427–1450.
[98] Amit Mangal. 2023. An Analytical Review of Contemporary AI-Driven Hiring Strategies in Professional Services. *ESP Journal of Engineering & Technology Advancements* 3, 3 (2023), 52–63.
[99] Robert W McGee. 2023. An AI Interview with Bruce Lee. *Available at SSRN 4643827* (2023).
[100] Qingfei Min, Haoye Sun, Xiaodi Wang, and Crystal Zhang. 2024. How do avatar characteristics affect applicants' interactional justice perceptions in artificial intelligence-based job interviews? *International Journal of Selection and Assessment* 32, 3 (2024), 442–450.
[101] Agata Mirowska and Laura Mesnet. 2022. Preferring the devil you know: Potential applicant reactions to artificial intelligence evaluation of interviews. *Human Resource Management Journal* 32, 2 (2022), 364–383.
[102] Dena F Mujtaba and Nihar R Mahapatra. 2019. Ethical considerations in AI-based recruitment. In *2019 IEEE International Symposium on Technology and Society (ISTAS)*. IEEE, 1–7.
[103] Nishad Nawaz. 2020. Artificial intelligence applications for face recognition in recruitment process. *Journal of Management Information and Decision Sciences* 23 (2020), 499–509.
[104] Josef Nguyen and Bonnie Ruberg. 2020. Challenges of designing consent: Consent mechanics in video games as models for interactive user agency. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. 1–13.
[105] Jakob Nielsen. 1994. Enhancing the explanatory power of usability heuristics. In *Proceedings of the SIGCHI conference on Human Factors in Computing Systems*. 152–158.
[106] Jeesun Oh, Wooseok Kim, Sungbae Kim, Hyeonjeong Im, and Sangsu Lee. 2024. Better to Ask Than Assume: Proactive Voice Assistants' Communication Strategies That Respect User Agency in a Smart Home Environment. In *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems*. 1–17.
[107] Ernesto Panadero and Anastasiya A Lipnevich. 2022. A review of feedback models and typologies: Towards an integrative model of feedback elements. *Educational Research Review* 35 (2022), 100416.
[108] Raja Parasuraman and Dietrich H Manzey. 2010. Complacency and bias in human use of automation: An attentional integration. *Human factors* 52, 3 (2010), 381–410.
[109] Hyanghee Park, Daehwan Ahn, Kartik Hosanagar, and Joonhwan Lee. 2021. Human-AI interaction in human resource management: Understanding why employees resist algorithmic evaluation at workplaces and how to mitigate burdens. In *Proceedings of the 2021 CHI conference on human factors in computing systems*. 1–15.
[110] Hyanghee Park, Daehwan Ahn, Kartik Hosanagar, and Joonhwan Lee. 2022. Designing fair AI in human resource management: Understanding tensions surrounding algorithmic evaluation and envisioning stakeholder-centered solutions. In *Proceedings of the 2022 CHI conference on human factors in computing systems*. 1–22.
[111] Veronica Pimenova, Sarah Fakhoury, Christian Bird, Margaret-Anne Storey, and Madeline Endres. 2025. Good Vibrations? A Qualitative Study of Co-Creation, Communication, Flow, and Trust in Vibe Coding. *arXiv preprint arXiv:2509.12491* (2025).
[112] Wan Ying Felicia Poh. 2015. Evaluating candidate performance and reaction in one-way video interviews. (2015).
[113] Postpone. 2024. List of Human Resources and Recruitment Subreddits. https://www.postpone.app/subreddit-lists/human-resources-and-recruitment

[114] Stanislav Pozdniakov, Roberto Martinez-Maldonado, Yi-Shan Tsai, Mutlu Cukurova, Tom Bartindale, Peter Chen, Harrison Marshall, Dan Richardson, and Dragan Gasevic. 2022. The question-driven dashboard: how can we design analytics interfaces aligned to teachers' inquiry?. In *LAK22: 12th international learning analytics and knowledge conference*. 175–185.

[115] Cassidy Pyle, Kat Roemmich, and Nazanin Andalibi. 2024. US Job-Seekers' Organizational Justice Perceptions of Emotion AI-Enabled Interviews. *Proceedings of the ACM on Human-Computer Interaction* 8, CSCW2 (2024), 1–42.

[116] Chuan Qin, Hengshu Zhu, Dazhong Shen, Ying Sun, Kaichun Yao, Peng Wang, and Hui Xiong. 2023. Automatic skill-oriented question generation and recommendation for intelligent job interviews. *ACM Transactions on Information Systems* 42, 1 (2023), 1–32.

[117] Manish Raghavan, Solon Barocas, Jon Kleinberg, and Karen Levy. 2020. Mitigating bias in algorithmic hiring: Evaluating claims and practices. In *Proceedings of the 2020 conference on fairness, accountability, and transparency*. 469–481.

[118] Leon Reicherts, Yvonne Rogers, Licia Capra, Ethan Wood, Tu Dinh Duong, and Neil Sebire. 2022. It's good to talk: A comparison of using voice versus screen-based interactions for agent-assisted tasks. *ACM Transactions on Computer-Human Interaction* 29, 3 (2022), 1–41.

[119] Lauren A Rivera. 2015. Go with your gut: Emotion and evaluation in job interviews. *American journal of sociology* 120, 5 (2015), 1339–1389.

[120] Kat Roemmich, Tillie Rosenberg, Serena Fan, and Nazanin Andalibi. 2023. Values in emotion artificial intelligence hiring services: Technosolutions to organizational problems. *Proceedings of the ACM on Human-Computer Interaction* 7, CSCW1 (2023), 1–28.

[121] Kat Roemmich, Florian Schaub, and Nazanin Andalibi. 2023. Emotion AI at work: Implications for workplace surveillance, emotional labor, and emotional privacy. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*. 1–20.

[122] Devjeet Roy, Xuchao Zhang, Rashi Bhave, Chetan Bansal, Pedro Las-Casas, Rodrigo Fonseca, and Saravan Rajmohan. 2024. Exploring llm-based agents for root cause analysis. In *Companion Proceedings of the 32nd ACM International Conference on the Foundations of Software Engineering*. 208–219.

[123] Richard M Ryan and Edward L Deci. 2000. Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American psychologist* 55, 1 (2000), 68.

[124] Md Nazmus Sakib, Ishika Tarin, Naga Manogna Rayasam, Manas Gaur, and Sanorita Dey. 2026. ReflectEd: Evaluating Reflection-Driven Learning in an AI-Assisted System. In *International Conference on Artificial Intelligence in Education*. Springer.

[125] Pedro Sanches, Axel Janson, Pavel Karpashevich, Camille Nadal, Chengcheng Qu, Claudia Daudén Roquet, Muhammad Umair, Charles Windlin, Gavin Doherty, Kristina Höök, et al. 2019. HCI and Affective Health: Taking stock of a decade of studies and charting future research directions. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. 1–17.

[126] Priya Sarkar and Cynthia CS Liem. 2024. " It's the most fair thing to do but it doesn't make any sense": Perceptions of Mathematical Fairness Notions by Hiring Professionals. *Proceedings of the ACM on Human-Computer Interaction* 8, CSCW1 (2024), 1–35.

[127] Benjamin Saunders, Julius Sim, Tom Kingstone, Shula Baker, Jackie Waterfield, Bernadette Bartlam, Heather Burroughs, and Clare Jinks. 2018. Saturation in qualitative research: exploring its conceptualization and operationalization. *Quality & quantity* 52 (2018), 1893–1907.

[128] Shreya Shankar, Rolando Garcia, Joseph M Hellerstein, and Aditya G Parameswaran. 2024. " We Have No Idea How Models will Behave in Production until Production": How Engineers Operationalize Machine Learning. *Proceedings of the ACM on Human-Computer Interaction* 8, CSCW1 (2024), 1–34.

[129] Monika Sieverding. 2009. 'Be Cool!': Emotional costs of hiding feelings in a job interview. *International Journal of Selection and Assessment* 17, 4 (2009), 391–401.

[130] Luke Stark and Jesse Hoey. 2021. The ethics of emotion in artificial intelligence systems. In *Proceedings of the 2021 ACM conference on fairness, accountability, and transparency*. 782–793.

[131] Ari Stillman and Jessica Witte. 2023. Measuring the Rise and Fall of a Subreddit using Text Analytics: The Life Course of r/Antiwork. In *2023 10th International Conference on Behavioural and Social Computing (BESC)*. IEEE, 1–7.

[132] Krishna Subramanian, Johannes Maas, Jan Borchers, and James Hollan. 2021. From detectables to inspectables: Understanding qualitative analysis of audiovisual data. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. 1–10.

[133] Hung-Yue Suen and Kuo-En Hung. 2023. Building trust in automatic video interviews using various AI interfaces: Tangibility, immediacy, and transparency. *Computers in Human Behavior* 143 (2023), 107713.

[134] Hung-Yue Suen and Kuo-En Hung. 2025. Comparing job applicant deception in asynchronous vs synchronous video interviews, with and without AI-assisted assessments. *Information Technology & People* 38, 2 (2025), 963–983.

[135] S Sumathi, D Kevin Harris, and AK Jeyanth. 2024. AI-Driven Interviewer: Enhancing Interview Experience Through Conversational AI. In *International Conference on Computational Intelligence in Data Science*. Springer, 180–196.

[136] Yuling Sun, Xiaojuan Ma, Silvia Lindtner, and Liang He. 2023. Care Workers' Wellbeing in Data-Driven Healthcare Workplace: Identity, Agency, and Social Justice. *Proceedings of the ACM on Human-Computer Interaction* 7, CSCW2 (2023), 1–29.

[137] Divy Thakkar, Neha Kumar, and Nithya Sambasivan. 2020. Towards an AI-powered future that works for vocational workers. In *proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. 1–13.

[138] Robert Thornberg, Kathy Charmaz, et al. 2014. Grounded theory and theoretical coding. *The SAGE handbook of qualitative data analysis* 5, 2014 (2014), 153–169.

[139] Sadia Nasrin Tisha, Md Nazmus Sakib, and Sanorita Dey. 2026. Competing or Collaborating? The Role of Hackathon Formats in Shaping Team Dynamics and Project Choices. In *Proceedings of the 57th ACM Technical Symposium on Computer Science Education V. 1*. 1061–1067.

[140] Edwin Torres, Amy M Gregory, and Cynthia Mejia. 2017. Interviews on demand: A case study of the implementation of asynchronous video interviews. *Journal of Hospitality & Tourism Cases* 5, 3 (2017), 23–37.

[141] Bianca Trinkenreich, Fabio Santos, and Klaas-Jan Stol. 2024. Predicting Attrition among Software Professionals: Antecedents and Consequences of Burnout and Engagement. *ACM Transactions on Software Engineering and Methodology* 33, 8 (2024), 1–45.

[142] William L Tullar. 1989. Relational control in the employment interview. *Journal of Applied Psychology* 74, 6 (1989), 971.

[143] April Tyack and Elisa D Mekler. 2024. Self-determination theory and hci games research: Unfulfilled promises and unquestioned paradigms. *ACM Transactions on Computer-Human Interaction* 31, 3 (2024), 1–74.

[144] Tom Van Nuenen, Jose Such, and Mark Cote. 2022. Intersectional experiences of unfair treatment caused by automated computational systems. *Proceedings of the ACM on Human-Computer Interaction* 6, CSCW2 (2022), 1–30.

[145] Qile Wang, Moath Erqsous, Kenneth E Barner, and Matthew Louis Mauriello. 2025. LATA: A Pilot Study on LLM-Assisted Thematic Analysis of Online Social Network Data Generation Experiences. *Proceedings of the ACM on Human-Computer Interaction* 9, 2 (2025), 1–28.

[146] Christiane Wenhart and Marc Hassenzahl. 2024. There is an "I" in "We": Relatedness Technologies Viewed Through the Lens of the Need for Autonomy. In *Extended Abstracts of the CHI Conference on Human Factors in Computing Systems*. 1–7.

[147] Jenny S Wesche and Andreas Sonderegger. 2021. Repelled at first sight? Expectations and intentions of job-seekers reading about AI selection in job advertisements. *Computers in human behavior* 125 (2021), 106931.

[148] Allison Woodruff, Renee Shelby, Patrick Gage Kelley, Steven Rousso-Schindler, Jamila Smith-Loud, and Lauren Wilcox. 2024. How knowledge workers think generative ai will (not) transform their industries. In *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems*. 1–26.

[149] Y Wayne Wu and Brian P Bailey. 2021. Better feedback from nicer people: Narrative empathy and ingroup framing improve feedback exchange. *Proceedings of the ACM on Human-Computer Interaction* 4, CSCW3 (2021), 1–20.

[150] Ziang Xiao, Xingdi Yuan, Q Vera Liao, Rania Abdelghani, and Pierre-Yves Oudeyer. 2023. Supporting qualitative analysis with large language models: Combining codebook with GPT-3 for deductive coding. In *Companion proceedings of the 28th international conference on intelligent user interfaces*. 75–78.

[151] Liying Xu, Yuyan Zhang, Feng Yu, Xiaojun Ding, and Jiahua Wu. 2024. Folk Beliefs of Artificial Intelligence and Robots. *International Journal of Social Robotics* (2024), 1–18.

[152] Lixiang Yan, Vanessa Echeverria, Gloria Milena Fernandez-Nieto, Yueqiao Jin, Zachari Swiecki, Linxuan Zhao, Dragan Gašević, and Roberto Martinez-Maldonado. 2024. Human-AI collaboration in thematic analysis using ChatGPT: A user study and design recommendations. In *Extended Abstracts of the CHI Conference on Human Factors in Computing Systems*. 1–7.

[153] Chi-Lan Yang, Alarith Uhde, Naomi Yamashita, and Hideaki Kuzuoka. 2025. Understanding and Supporting Peer Review Using AI-reframed Positive Summary. In *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems*. 1–16.

[154] Xi Yang and Marco Aurisicchio. 2021. Designing conversational agents: A self-determination theory approach. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. 1–16.

[155] Xiliu Yang, Prasanth Sasikumar, Felix Amtsberg, Achim Menges, Michael Sedlmair, and Suranga Nanayakkara. 2025. Who is in Control? Understanding User Agency in AR-assisted Construction Assembly. In *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems*. 1–15.

[156] Chen Zhang and Hao Wang. 2018. Resumevis: A visual analytics system to discover semantic information in semi-structured resume data. *ACM Transactions on Intelligent Systems and Technology (TIST)* 10, 1 (2018), 1–25.

[157] Annuska Zolyomi and Jaime Snyder. 2024. An Emotion Translator: Speculative Design By Neurodiverse Dyads. In *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems*. 1–18.

A Appendix

A.1 Keywords used to filter Reddit comments

AI interviewer [94], AI interviews [99], AI-driven interview [135], automated interview [86], AI-based video interview [133], Algorithmic hiring [117], AI tools in recruitment [63], AI-based interview [90], AI-powered interviews [18, 81], AI-driven hiring [98], automation in recruitment [56], virtual interview [72], asynchronous video interview [96], video based interview [5], one way video interview [112]

A.2 Interview Questionnaire

- Can you share your overall experience with AI interviewers?
- Did the AI interview tool meet your expectations? If not, what was missing or different from what you had anticipated?
- Did you receive any guidance or instructions from the organization on how to prepare for the AI interview?
- How confident did you feel while responding to the AI interviewer? Were there any particular reasons for feeling confident or unsure?
- Did you face any technical difficulties, such as internet issues, lag, or audio/video problems? How did you manage or resolve them?
- Were there any issues with how long you had to respond or how your responses were timed?
- Were you able to review or re-record your answers, or did you feel pressured to get it right in one take? Would the ability to edit your responses have changed your experience?
- What is your impression of how AI systems evaluate candidates? Do you have any thoughts or assumptions about how the decisions are made?
- Did you have any concerns about potential bias in the system,for example, based on your accent, appearance, neurotype, or background?
- Before the interview, were you aware of which AI tool or platform you'd be using?
- Did the interview questions feel personalized or adaptive to you, or did they seem generic and one-size-fits-all?
- Some people say the questions in AI interviews are short and lacking in depth. Did you experience something similar?
- While recording your responses, did you encounter any challenges (e.g., time pressure, speaking naturally, staying focused)? How did you deal with them?
- Did the AI interviewer simulate human-like behavior in any way (e.g., facial cues, verbal responses, timing)? If so, how did that affect your experience?
- Were you expected to maintain specific facial expressions, look into the camera, or behave in a certain way during the recording?
- From your perspective, what was one important element missing from the AI interview process?
- How transparent did the AI interview feel to you? Were you informed about how your responses would be processed or evaluated?
- Was the interview recorded, and were you notified whether your facial expressions, voice, or other biometric data would be analyzed?
- Did the overall process feel respectful and fair, or did it feel impersonal or dismissive in any way?
- Were there any parts of the interview that felt inaccessible or difficult to navigate due to your personal or neurodivergent needs?
- Overall, were you satisfied with the AI interview experience? What would you change or improve?
- Compared to a human interviewer, did you feel more or less comfortable facing the AI system? Why?

A.3 Few-Shot Prompting to Identify Relevant Reddit Data

System Prompt: You are a helpful assistant trained to identify whether a Reddit data point describes a person's experience with AI-based asynchronous interviews. These interviews are typically one-way, where applicants face AI-based interview without a human interviewer.

User Prompt: Below are several Reddit data points. For each entry, respond with either "Relevant" or "Irrelevant", followed by a short justification explaining your decision. A relevant data point should mention or strongly imply personal experience (positive, negative or neutral) with asynchronous interviews that may or may not involve AI tools, such as HireVue or similar systems. Followings are a few examples:

Examples:
Comment 1:

"I had to stare at my webcam and talk to a bot for 20 minutes. It was exhausting and felt like talking into the void."
Response:
Relevant – This comment clearly refers to a one-way asynchronous interview experience.

Comment 2:
"My friend told me these AI interviews are just a scam. I haven't done one myself."
Response:
Irrelevant – The speaker did not participate in the interview themselves.

Comment 3:
"I've used Zoom for a few interviews. Nothing fancy."
Response:
Irrelevant – This refers to synchronous, human-led interviews.

A.4 Behavioral Questions

These following questions are collected from Indeed [80].

Used in Response Variants:
1. Give me an example of a time you had a conflict with a team member. How did you handle it?
2. Tell me about a time you made a mistake at work. How did you resolve the problem, and what did you learn from your mistake?
3. Describe an occasion when you had to use time management skills to complete a task or meet a deadline. How did you do it?

Used in Feedback Variants:
1. Describe an occasion when you failed at a task. What did you learn from it?
2. Tell me about a time you took the initiative in your career. What was your motivation for doing so?
3. Describe a time when you used your leadership skills to motivate your team or colleagues.

A.5 Few-Shot Prompting To Generate Feedback

A.5.1 Motivational Feedback. You are a supportive mentor giving motivational feedback to someone who is responding to behavioral interview responses. Your job is to:

- Encourage their effort
- Reinforce their self-reflection
- Avoid any kind of judgment, critique, or correction
- Be warm, affirming, and growth-oriented
- Keep the tone friendly and inspiring
- Write in 1–3 natural, uplifting sentences

Use the examples below as guidance:
Example 1
Q1: Give me an example of a time you had a conflict with a team member. How did you handle it?
Response: "I had a disagreement with a teammate about how to divide tasks. We had different views on priorities, but eventually we talked it through and came to a compromise."
Motivational Feedback: *"Thank you for sharing this experience. It takes courage to reflect on conflict. You're clearly engaging with the story and showing awareness, and that's an important step forward. Keep practicing and trusting your instincts, you're doing great."*
Example 2
Q2: Tell me about a time you made a mistake at work. How did you resolve the problem, and what did you learn from your mistake?
Response: "I forgot to send out an important email to a client. Once I realized, I contacted them right away to apologize and sent the information. It made me more careful with deadlines in the future."
Motivational Feedback: *"That's a brave reflection, thanks for being open about it. It's clear you care about doing well and learning from experiences. Keep it up, this kind of insight shows real growth."*
Example 3
Q3: Describe an occasion when you had to use time management skills to complete a task or meet a deadline. How did you do it?

Response: "I had to juggle multiple deadlines during finals week. I made a daily schedule, blocked out time for each task, and avoided distractions. I managed to submit everything on time."

Motivational Feedback: *"Well done ! You're showing commitment and intention, and that's something to be proud of. Taking time to reflect on how you managed it is a sign of strong self-awareness. Keep up the great work!"*

A.5.2 Evaluative Feedback. You are a thoughtful interviewer giving evaluative feedback on behavioral responses. For each response you read, write feedback that is:

- **Friendly and natural in tone**
- **Evaluative and helpful**
- **Framed around the STAR method (Situation, Task, Action, Result)**
- **Suggestive, not judgmental, ike a helpful mentor or coach**
- **Written in 2–4 sentences**

Use these examples as a guide:

Example 1

Q1: Describe an occasion when you failed at a task. What did you learn from it?

Response: "I was leading a small marketing campaign and forgot to confirm our ad schedule with the publishing team. We missed our planned launch window. I took responsibility and personally contacted the team to find an alternate slot, which we secured the following week. Since then, I always double-check timelines and confirmations."

Feedback: *"Thanks for sharing that candidly. You explained the situation and your role well, and your ownership is commendable. To strengthen this further, you could expand on what concrete strategies you now use to avoid similar mistakes,this helps make your learning more visible."*

Example 2

Q2: Tell me about a time you took the initiative in your career. What was your motivation for doing so?

Response: "During a slow quarter, I suggested we repurpose existing content into short-form videos for social media. I drafted scripts and coordinated with the design team to get them out quickly. Engagement went up by 30%. I wasn't told to do it, but I knew we needed to boost visibility."

Feedback: *"This is a solid example of taking initiative with a measurable impact. To deepen your answer, it might help to explain how you identified the opportunity in the first place and what feedback you received afterward. That kind of reflection highlights your awareness and influence."*

Example 3

Q3: Describe a time when you used your leadership skills to motivate your team or colleagues.

Response: "We were behind on a group project. I called a team huddle, broke down tasks into manageable parts, and created a progress board everyone could update. People became more engaged and we completed it on time."

Feedback: *"This shows thoughtful coordination and problem-solving under pressure. You might enhance it by sharing how your leadership approach impacted the team dynamic or whether it led to lasting change. Specific feedback from team members could also enrich the story."*

A.6 Tool Names Found in Our Study

HireVue, Sapia AI, Talview, Vervoe, XOR, Willo, SparkHire, Clovers AI, Skillora AI, Paradox AI, Eightfold AI, Echotalent AI, Vhire, MyInterview, Quinncia

Received May 13, 2025; revised January 13, 2026; accepted March 17, 2026