Journal of Computer-Mediated Communication, 2023, **28(4)**, zmad028
https://doi.org/10.1093/jcmc/zmad028
Full-length Research Article
Special Issue: Technology and the Future of Work

International Communication Association
OXFORD

Speech production under uncertainty: how do job applicants experience and communicate with an AI interviewer?

Bingjie Liu [1,*], Lewen Wei [2], Mu Wu [1], Tianyi Luo[3]

[1]Department of Communication Studies, California State University, Los Angeles, CA, USA
[2]School of the Arts and Media, University of New South Wales, Sydney, Australia
[3]Amazon Science, Amazon.com, Inc., Palo Alto, CA, USA

*Corresponding author: Bingjie Liu. Email: bliu3@calstatela.edu

Abstract

Theories and research in human–machine communication (HMC) suggest that machines, when replacing humans as communication partners, change the processes and outcomes of communication. With artificial intelligence (AI) increasingly used to interview and evaluate job applicants, employers should consider the effects of AI on applicants' psychology and performance during AI-based interviews. This study examined job applicants' experience and speech fluency when evaluated by AI. In a three-condition between-subjects experiment ($N = 134$), college students had an online mock job interview under the impression that their performance would be evaluated by a human recruiter, an AI system, or an AI system with a humanlike interface. Participants reported higher uncertainty and lower social presence and had a higher articulation rate in the AI-evaluation condition than in the human-evaluation condition. Through lowering social presence, AI evaluation increased speech rate and reduced silent pauses. Findings inform theories of HMC and practices of automated recruitment and professional training.

Lay Summary

Recently, major employers have adopted artificial intelligence (AI) systems to interview and evaluate job applicants to inform hiring decisions. From the perspective of job applicants, conversing with and being evaluated by an unfamiliar, nonhuman agent may induce additional uncertainty, and lead them to feel and communicate differently than having an interview with a person. Evaluations based on such distorted responses from job candidates could be misleading. Against this backdrop, we examined college students' experience and performance during mock job interviews when they were told that an AI system (vs. a human recruiter) would evaluate their performance. Participants reported experiencing higher uncertainty when evaluated by AI than by the human recruiter. They also reported having less social presence (i.e., the sense of being together with a person) during AI-based interviews, which led them to speak faster and pause less frequently. Overall, results suggest that the differences between the outcomes of AI evaluation and human evaluation can be nuanced. We suggest designers of AI systems consider such differences to evaluate job candidates more accurately and recommend job applicants overcome uncertainty about AI and compensate for the reduced social presence by acting more engaged during AI-based interviews to succeed.

Keywords: AI decision-making, job interview, uncertainty, social presence, speech production

Job interviews are high-stakes settings for job applicants when they are judged and expected to showcase their abilities to the best during a brief interaction with the interviewers. Communicating fluently is therefore important during job interviews, as it not only demonstrates one's communication skill, a highly demanded quality in the workplace (Robles, 2012), but also exemplifies one's expertise and competence (Brennan & Williams, 1995). One recent innovation in interview methods, however, may debilitate job candidates' speech production during job interviews.

Increasingly, artificial intelligence (AI), defined as machines that can do tasks that require human intelligence (e.g., decision making, Luger & Stubblefield, 1997), is deployed to interview and evaluate job applicants. During a typical AI-based interview, job applicants talk through their computer cameras while the AI system collects and analyzes their responses, both verbal and nonverbal (e.g., facial movements), based on which the AI system rates their employability to inform hiring decisions (Harwell, 2019). Such AI-based interview systems have already been adopted by many major employers in their recruitment processes (Jaser et al., 2022).

Interacting with machines, however, alters the original communication context of human–human interaction and may lead individuals to communicate differently (Gambino & Liu, 2022). Previous research found that job applicants often reported feeling uncertain about the evaluation process and criteria of those AI-based interview systems (Mirowska & Mesnet, 2022; Schick & Fischer, 2021). According to Berger's plan-based theory of strategic communication, when someone has high uncertainty about their conversational partner (e.g., unsure about their states and abilities), planning becomes cognitively demanding and can debilitate speech production (Berger, 1995, 1997) and decrease speech fluency (Berger & Jordan, 1992). Speech fluency, defined as the extent to which speeches are smooth, fluid, and free of speech errors (Greene & Lindsey, 1989), has been found to influence the evaluation of job applicants made by both human recruiters (e.g., Brosy et al., 2016; Martín-Raugh et al., 2023) and algorithms trained with human ratings as the ground truth (e.g., Naim et al., 2018). With greater fluency usually predicting more positive evaluation, should job applicants deliver less fluent speech due to their uncertainty about AI,

Associate Editors: Nancy Baym and Nicole Ellison
Received: 2 October 2022. **Revised:** 15 May 2023. **Accepted:** 18 May 2023

applying the evaluation standards derived from human-led interviews indiscriminately could lead to incorrect, unjust rejection. To avoid such mistakes, it is imperative to calibrate and then account for the net effects of adopting an AI-based interview system on job applicants' psychological experience and communication when employers make hiring decisions.

So far, research has examined individuals' reactions to AI-based interviews relative to human-led interviews mostly with vignette-based designs, where participants saw a video (e.g., Langer et al., 2019) or a textual description of AI-based hiring procedures (e.g., Gonzalez et al., 2022; Köchling et al., 2022; Mirowska & Mesnet, 2022; Schick & Fischer, 2021) and were asked to imagine how they would feel in such situations. Nevertheless, little research had participants go through the interview process to examine their experience and speech performance. As individuals' operant psychological processes during real-time interaction with AI may differ from those reported in imaginative settings, it remains unclear whether and how AI evaluation affects job applicants during job interviews. The present study aims to fill this gap by having participants go through mock job interviews to investigate their experience and performance during AI-based interviews in terms of their uncertainty, social presence, and speech fluency.

Literature review

AI-based interviews in hiring: agency shift and potential outcomes

In a traditional hiring process, the hiring team screens job applicants' resumes, conducts interviews with them, remotely or onsite, evaluates their performance, determines their aptitude for a given job, and then finally decides whether an offer should be extended. In such cases, the decision-making agency fully remains with the human recruiters. Recently, many employers have started using AI to inform some or all of those steps of recruitment (Langer et al., 2019; Naim et al., 2018). In such AI-informed hiring, although human recruiters can still exercise some decision-making agency through designing interview questions or setting certain parameters (e.g., the time allowed for each answer), they neither make decisions based on their perceptions of a candidate nor do they set explicit rules for the AI system to follow. Instead, modern AI systems use machine learning "to define or modify decision-making rules autonomously" (Mittelstadt et al., 2016, p. 3), for example, by identifying patterns of how certain features of one's answers (e.g., pauses) correspond to their competence (Adepu et al., 2020). As such, AI's decision rules are largely data-driven, not predetermined by humans, and often uninterpretable to humans (Liu, 2021). Hence, with AI used for candidate evaluation, informing, and even making the final hiring decisions, the agency of decision making in recruitment shifts from humans to machines.

Except for very few studies that found positive perceptions of AI evaluation (for its higher consistency, Acikgoz et al., 2020; Kim & Heo, 2021), research overall found negative sentiments toward such an agency shift among both job applicants and employers (see a review, Langer & Landers, 2021). Two recurring themes emerging in qualitative studies are job applicants' concerns about the opaqueness of AI's evaluation method and the perception of lacking control (Kim & Heo, 2021; Mirowska & Mesnet, 2022), both of which seem to be symptoms of uncertainty about the AI systems that evaluate them.

Uncertainty, according to theories and research in message production, can complicate planning and disrupt speech production (Berger, 1997). Hence, adopting AI interview systems may compromise job applicants' speech fluency during job interviews. While the content of job applicants' answers matters, research has found vocal features including speech fluency influence evaluation and hiring decisions above and beyond the content effect (Naim et al., 2018). In interpersonal communication, speech fluency is commonly used by listeners as an indicator of the speakers' knowledge, metacognitive states (Brennan & Williams, 1995), intelligence and competence (Breil et al., 2021), and personalities such as conscientiousness and openness (Breil et al., 2021; Martín-Raugh et al., 2023). In both human-led (Brosy et al., 2016; Martín-Raugh et al., 2023) and AI-based hiring (Naim et al., 2018), fluent speech leads to more positive evaluation and a greater chance of employment as it signals higher competence. Should job applicants respond less fluently because of their uncertainty about AI, the undermined speech performance could be merely an artifact of interacting with an AI system (as opposed to a human). Such "AI effects," if misread as a sign of one's incompetence, will lead to inaccurate and potentially discriminating hiring decisions, with marginalized groups such as those with low AI literacy and those with speech-related disabilities more severely harmed.

Given the importance of speech fluency in evaluating job candidates, this study focuses on the effects of AI on interviewees' speech fluency and pertinent psychological mechanisms. Following the suggestion from theories and research in speech fluency (Tavakoli & Wright, 2020), we considered both (a) speed fluency, which measures the flow and continuity of speech, and (b) breakdown fluency, i.e., pauses that break down the speech flow. We adopted the commonly used measures within each category, with speech rate and articulation rate measuring speed fluency, and with onset latency (Knowlton & Berger, 1997), silent pauses, and filled pauses measuring breakdown fluency (Knobloch, 2006; Pickard & Roster, 2020). Notably, while both articulation rate and speech rate are considered measures of speed fluency, speech rate is calculated as the number of syllables divided by the total speech performance time, whereas for articulation rate, the denominator excludes silent pause intervals. Therefore, articulation rate is a "pure" measure of speed, and speech rate is a "composite" measure of both speed and breakdown (Tavakoli & Wright, 2020, p. 48).

Predicting the effects of AI evaluation on speech fluency, however, is not straightforward. As Gambino and Liu (2022) suggested, processes and outcomes of human–machine communication (HMC) depend on the contexts, across which individuals' goals vary, so do the operant psychological processes. Besides greater uncertainty, job applicants may experience less social presence with machines (vs. humans), which could instead facilitate speech production in certain contexts (e.g., when disclosing sensitive information, Pickard & Roster, 2020). It is still unclear which process is more dominant in the context of job interviews. Drawing on existing theories and research, below we discuss these two potentially competing mechanisms—uncertainty and social presence—and discuss how AI evaluation (vs. human evaluation) might affect speech fluency during job interviews via each.

Effects of machine agency on job applicants' experience: uncertainty

In interpersonal communication, uncertainty is defined as one's inability to explain and predict their interactant's behaviors and internal states such as their beliefs and attitudes (Berger, 1997). For individuals to communicate competently, form communication plans, and achieve goals through communication, it is imperative for them to reduce uncertainty, such as by obtaining information about their interactant (Berger, 1997). For decision-making AI, uncertainty about it often arises from individuals' lack of knowledge about its decision-making processes (Liu, 2021). During AI-based interviews, job applicants may experience high uncertainty because without knowing the AI system's evaluation methods, they can hardly explain or predict its evaluations (Kim & Heo, 2021; Langer et al., 2019; Schick & Fischer, 2021).

Moreover, reducing uncertainty about AI interviewers is not easy in practice. Those systems are often proprietary and do not reveal how and why evaluations are made (Whittaker et al., 2019). With human interviewers, however, the norms and conventions that humans share (Miller & Steinberg, 1975) can serve as a mental model, i.e., a structured mental representation of the state of affairs in the world (Johnson-Laird, 1980), allowing interviewees to explain and predict the behaviors of their human interactants (Rouse & Morris, 1986), so as to reduce interpersonal uncertainty.

Yet, with machine interactants, the mental model of a generic person is no longer applicable. Without the guidance of anthropocentric knowledge, individuals might experience higher uncertainty about machines than humans. For example, in human–robot interaction, individuals reported having higher uncertainty about social robots than humans (Edwards et al., 2016). When knowing that it would be a robot (vs. a human) evaluating their public speaking performance, students showed greater apprehension which could result from their higher uncertainty about the robot (Edwards et al., 2018). Even those who regarded AI as more objective than humans still preferred human-led interviews, because they preferred "the devil they know"; as put by one participant, "I know how humans think, even though I don't know the human[s] themselves. I have an idea how they think more or less. But the AI, I don't know. [T]his uncertainty would make me feel extra nervous" (Mirowska & Mesnet, 2022, p. 12). Based on these research findings, we propose the following hypothesis.

> **H1.** AI evaluation (vs. human evaluation) induces higher uncertainty during job interviews.

Effects of machine agency on speech production: the role of uncertainty

Job interviews are episodes of goal-directed strategic communication when applicants strive to achieve the goal of being hired through communication with the interviewers. To achieve goals, individuals usually form communication plans—cognitive representations of action sequences—to guide their speech (Berger, 1995; Berger & Jordan, 1992). According to Berger's plan-based theory of strategic communication (Berger, 1995, 1997), effective planning is contingent on one's knowledge about their communication partner's states and capabilities and their own ability to adjust accordingly. When individuals are unsure about their communication partner's states and abilities (i.e., having high uncertainty), it is difficult to devise action sequences (e.g., deciding what to say first), to predict the outcomes of their speech, and to adjust accordingly (Berger & Bradac, 1982). Berger also argued that under uncertainty, planning becomes more bottom-up, opportunistic, and conscious (Berger, 1997), and compared with top-down planning, bottom-up planning takes more time and energy to consciously develop, execute, and modify (Berger, 1988). In other words, planning requires more resources when there is high uncertainty.

As individuals become more preoccupied with planning under higher uncertainty, a corollary follows that they will have higher rates of encoding disruptions in communicative actions (Berger, 1997), i.e., their speech fluency will decrease (Berger & Jordan, 1992). For example, in dating contexts, when individuals had higher uncertainty about their partners, they found it more difficult to have a conversation (Knobloch & Solomon, 2005). When individuals produced messages of a date request to their partners, their relational uncertainty was negatively associated with their speech fluency and message effectiveness (Knobloch, 2006). In sum, lacking information about one's communication partner may compromise individuals' ability to plan their speech and thereby disrupt speech production.

So far, research is still limited in understanding the effects of machine agency on individuals' message planning and production, but a few studies that examined related concepts found negative effects of AI on strategic communication. For example, Langer et al. (2020a) had participants answer job interview questions via voice recordings and told them that their answers would be analyzed and evaluated either by a person or by a computer; participants in the computer-evaluation condition reported having fewer opportunities of impression management and provided shorter answers. Similarly, vignette-based studies also found individuals perceive less chance to perform (Acikgoz et al., 2020; Köchling et al. 2022), lower self-efficacy (Gonzalez et al., 2022; Langer et al., 2019), and lower control (Gonzalez et al., 2022) when evaluated by AI during job interviews. Taken together, AI evaluation seems to limit job applicants' ability to communicate strategically during job interviews. Therefore, we anticipate individuals becoming less competent in communicating with machines (than with humans) due to heightened uncertainty and producing less fluent speech.

> **H2.** AI evaluation (vs. human evaluation) has a negative indirect effect on individuals' speech fluency by inducing higher uncertainty.

Effects of machine agency on speech production: the role of social presence

Nevertheless, previous research suggests that machine agency may instead facilitate speech production by triggering less social presence, i.e., the perception of being together with another person (Blascovich, 2002; Edwards et al., 2021). According to Blascovich's (2002) Threshold Model of Social Influence, humans induce greater social presence than artificial agents in social interactions, and greater social presence results in greater social influence (meta-analysis, Fox et al., 2015). For example, in immersive virtual environments, the presence of human-controlled avatars, but not computer-controlled agents, caused social facilitation (Hoyt et al., 2003), i.e., having an inhibition effect on novel tasks (Zajonc, 1965). Consistent with this perspective, Gambino and Liu (2022) argued that conversing with humans may trigger greater social concerns in terms of impression and relationship management that are less applicable to HMC. With fewer concerns in HMC, individuals will have a less complex goal structure (Dillard et al., 1989), which demands less complex planning, and should produce more fluent speech (Knowlton & Berger, 1997). For example, Pickard and Roster (2020) examined individuals' answers to sensitive questions (e.g., "What do you feel most guilty about in your life?") asked by a human (face-to-face), a computer with a human avatar, or a voice-only computer. They found that participants disclosed more sensitive information (as rated by independent judges) and spoke more fluently to the voice-only computer than to the human, with the computer-with-human-avatar condition ranked in the middle on those measures. Thus, we hypothesize that AI evaluation in job interviews, by reducing social presence, may facilitate speech production.

H3. AI evaluation (vs. human evaluation) induces less social presence during job interviews.
H4. AI evaluation (vs. human evaluation) has a positive indirect effect on individuals' speech fluency by reducing social presence.

Given the two potential competing mechanisms underlying the effects of machine agency on speech fluency, we propose the following research question.

RQ1. What are the effects of AI evaluation (vs. human evaluation) on individuals' speech fluency during job interviews?

Anthropomorphic cues as a remedy?

Besides examining the potentially negative effects of AI evaluation on speech production, we also explored potential solutions. Can an AI system be designed in ways that reduce uncertainty so as to mitigate its negative effects on planning and speech production? The three-factor theory of anthropomorphism suggests that a humanlike interface may help reduce individuals' uncertainty about machines by rendering anthropocentric knowledge accessible (Epley et al., 2007). For example, compared with a machine-like robot, a humanlike robot triggered a richer and more anthropomorphic mental model among individuals (Kiesler & Goetz, 2002), and individuals self-reported having more knowledge about the humanlike robot (Ososky et al., 2013). Also, the more participants felt a robot was humanlike, the less uncertainty they had about it (Edwards et al., 2016). Therefore, adding a humanlike avatar to the AI system, a common representation during virtual interviews, may help reduce uncertainty, guide planning, and improve speech fluency for job applicants.

Alternatively, adding anthropomorphic cues may also increase social presence (Blascovich, 2002; Oh et al., 2018), thereby raising additional social concerns, complicating goal structures, and debilitating speech production. We raise the following hypothesis and research question.

H5. Compared with an AI system without visual anthropomorphic cues, adding a humanlike avatar (a) reduces uncertainty and (b) increases social presence.
RQ2. What are the effects of using a humanlike avatar (vs. no visual anthropomorphic cues) on individuals' speech fluency in AI-based interviews?

Method

Overview

We conducted a three-condition between-subjects online experiment with college students in the United States. Participants first answered a questionnaire on *Qualtrics.com* asking about their interview experience, trait anxiety, and sense of uniqueness. Participants were then prompted to think about and type down a job position that they wanted to apply for and to treat the following mock interview as the actual interview for that job. Participants were randomly assigned to one of three conditions; they were told that their performance would be reviewed and evaluated by (a) a human resource expert, (b) an AI system without visual anthropomorphic cues (referred to as "plain AI" hereafter), or (c) an AI system with a humanlike avatar to determine their job readiness. During the mock interview, six common, generic questions were asked. We used *Pipe* (https://addpipe.com), an application embedded in Qualtrics, to record participants' answers. After the interview, participants answered questions on their experience of the interview, demographic information, and their knowledge and belief about AI.

Participants

Participants were college students from public universities in the United States and were rewarded with extra credit. Among the 173 complete cases, three were duplicates, and 36 failed the manipulation check. Therefore, the final data set consisted of 134 valid responses.

Among the 134 participants, 39 self-reported as male, 90 as female, and 5 as non-binary. Their age ranged from 18 to 44 years ($M = 24.21$ years, $SD = 5.91$). Sixty-eight participants self-identified as Hispanic/Latino/a, 22 as White, 18 as Asian, 14 as Black or African American, 1 as Native Hawaiian or Pacific Islander, 3 as Other, and 8 as Multiracial.

Experiment materials and manipulation

Interview questions

Considering participants' different backgrounds and career plans, we used six standard questions that are common in job interviews (Langer et al., 2020a, b; Suen et al., 2019), such as "What is your greatest weakness, and how do you plan to overcome, or work with it?" In the human and plain-AI conditions, interview questions were presented in textual form; in the humanlike-AI condition, interview questions were presented in videos featuring a speaking avatar. For each question, participants had a maximum of 3 minutes to answer, a commonly used time limit in automated job interviews (Langer et al., 2017, 2020a). After seeing each question, participants had 10 seconds to prepare before proceeding to the next page with the recorder embedded, where recording started immediately, with a 3-minutes countdown displayed in the recorder window. After 3 minutes, recording would automatically stop, and the next question would show up. If participants finished answering within 3 minutes, they manually stopped the recording and proceeded.

Manipulation

The agency manipulation (i.e., instructions on who/what would review and evaluate participants' performance) appeared both before and during the mock interview so that participants were repeatedly informed that their answers would be "reviewed and evaluated by a human resource expert" or "by an AI system."

In the humanlike-AI condition, we used two visual representations of the job interviewer selected from a pretest. In the pretest, we first selected eight avatars produced by *Synthesia* (https://www.synthesia.io/), an online platform using AI technology to create videos based on professional-looking, real human actors for professional use (Synthesia, 2022). The eight avatars, who are all dressed in professional outfits, were chosen for their apparent naturalness in speaking and facial movements, and they varied in physical appearance embodied by sex (i.e., Female and Male) and race/ethnicity (i.e., Asian, Black, Latino(a)/Hispanic, and White). We then recruited 74 college students (M_{age} = 28.70 years; Female = 67.1%) through *CloudResearch*, an online platform providing affordable access to survey-takers on Amazon's Mechanical Turk (CloudResearch, 2022); each participant saw two videos of one of the eight avatars raising interview questions and evaluated the avatar's humanlikeness, eeriness (Ho & MacDorman, 2010), and commonality. We chose the two avatars of the highest humanlikeness, highest commonality, and lowest eeriness—one White female avatar (Mallory) and one Black female avatar (Evelyn)—to create the interview videos used in the main study (Figure 1). We adopted these two avatars instead of all eight to avoid introducing additional within-group variance and to ensure construct validity (i.e., humanlikeness) and ecological validity (i.e., companies are unlikely to adopt avatars that induce eeriness).

Measurements

If not specified otherwise, all measures were on a 7-point scale. Details of all measures are reported in the Supplementary Material.

Manipulation check

After the interview, manipulation success was checked by asking participants to answer "According to the instruction you received, how would your answers in this interview be reviewed and evaluated?" by choosing one of the three choices: "By an AI system," "By a human resource expert," and "It was not mentioned." Thirty-six participants failed this manipulation check and were removed from the dataset.

Measures of individual differences

As suggested by previous research, individuals' reactions to AI evaluation may be influenced by their knowledge about AI (Araujo et al., 2020), belief in AI's objectivity (Kim & Heo, 2021; Liu, 2021), communication anxiety (Constantin et al., 2021), experience with video interviews (Langer et al., 2019; Suen et al. 2019), and perceived uniqueness of themselves (Longoni et al., 2019). We, therefore, included these individual differences as covariates.

Sense of uniqueness. We measured one's sense of uniqueness with three items adapted from the Personal Sense of Uniqueness Scale (Şimşek & Yalınçetin, 2010), such as "I feel

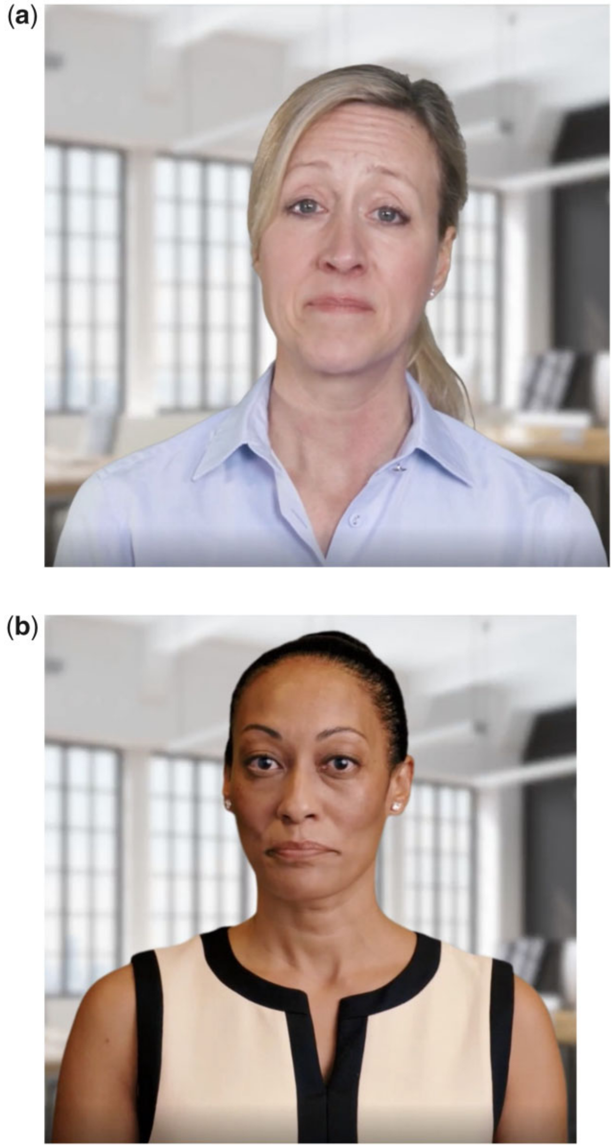

Figure 1. Avatars of interviewer in the humanlike-AI condition: (a) Mallory; (b) Evelyn.

that some of my characteristics are completely unique to me." Cronbach's $\alpha = 0.872$, $M = 5.17$, $SD = 1.23$.

Video interview experience. We asked participants "How experienced are you with online asynchronous video job interviews?" with two items, "not familiar at all/extremely familiar" and "not experienced at all/extremely experienced." Cronbach's $\alpha = 0.971$, $M = 2.71$, $SD = 1.80$.

Communication anxiety. Dispositional communication anxiety was measured with four items adapted from the Communication Anxiety subscale (McCarthy & Goffin, 2004), such as "I can be so apprehensive that I am unable to express my thoughts clearly." Cronbach's $\alpha = 0.806$, $M = 3.74$, $SD = 1.36$.

Belief in AI objectivity. Belief in AI objectivity was measured with four items (Liu, 2021), such as "If a decision is made by AI, it must be accurate." Cronbach's $\alpha = 0.868$, $M = 3.73$, $SD = 1.23$.

Self-perceived knowledge about AI. Self-perceived knowledge about AI was measured by asking participants to rate their knowledge in three areas: AI, computer programming, and machine learning (Araujo et al., 2020). Cronbach's α = 0.919, $M = 3.24$, $SD = 1.71$.

Measures of dependent variables

Certainty. Certainty was measured with five items assessing one's ability to predict and explain a decision-making agent (Clatterbuck, 1979; Liu, 2021), such as "I felt I was able to predict the human resource expert's/AI system's evaluation of me." Cronbach's $\alpha = 0.837$, $M = 3.83$, $SD = 1.26$.

Social presence. Social presence was measured with three items (Lee et al., 2006), such as "During the job interview, how much did you feel that you were in the company of an interviewer?" Cronbach's $\alpha = 0.889$, $M = 3.85$, $SD = 1.77$.

Speech fluency. We used *Pratt*, a software program that analyzes phonetics, and a validated algorithm written in Pratt script to measure (a) speed fluency and (b) breakdown fluency of participants' responses to all six questions (De Jong et al., 2021). Speed fluency measures included speech rate ($M = 3.23$ syllables/second [syll/s], SD = 0.48) and articulation rate ($M = 4.49$ syll/s, SD = 0.38) (Tavakoli & Wright, 2020). Breakdown measures included onset latency ($M = 1.46$ s, SD = 0.55), the average frequency of silent pauses (M = 0.48 pauses/s, SD = 0.17), and the average frequency of filled pauses (M = 0.14 pauses/s, SD = 0.09). Details of their conceptual and operational definitions are in Table 1.

Bivariate correlations of dependent variables are reported in the Supplementary Material.

Results

After confirming that neither of the two avatars was perceived as eerie (Mallory: $M = 3.38$, $SD = 1.53$; Evelyn: $M = 3.06$, $SD = 1.46$), we combined participants in those two humanlike-AI conditions and analyzed them as one group. Descriptive statistics by avatar are reported in the Supplementary Material.

Hypotheses testing

We test the hypotheses with a set of analysis of covariance (ANCOVA), with five individual differences as covariates, namely previous video interview experience, belief in AI objectivity, self-perceived knowledge about AI, communication anxiety, and sense of uniqueness. Power analysis with G*Power (Faul et al., 2009) suggests that with a sample of $N = 134$, α = 0.05, and five covariates, the power to detect any medium-size ($f = 0.25$) main effect in ANCOVA is 0.73, and the power to detect any small-size ($f = 0.10$) main effect is only 0.16. Hence, in addition to reporting point estimates and 95% confidence intervals (CIs), we also reported 90% CI to detect small effects for future research to validate, when applicable.

Table 2 summarizes the ANCOVA results. When examining the effects of AI evaluation, in support of H1 and H3, we found AI evaluation induced lower certainty ($M_{\text{difference}}$ = −0.61, p = .014, 95% CI [−1.10 to −0.13], Cohen's $d = -0.523$) and less social presence ($M_{\text{difference}} = -0.84$, p = .020, 95% CI [−1.55 to −0.14], Cohen's $d = -0.505$) than human evaluation. In answering RQ1, we found no significant effects of AI evaluation on any speed-related or breakdown-related measures. With a small- to medium-size effect that was close to significance, we found AI evaluation led to a higher articulation rate than human evaluation ($M_{\text{difference}} = 0.16$, $p = .054$, 95% CI [−0.003 to 0.31], 90% CI [0.02 to 0.29], Cohen's $d = 0.439$).

When examining the effects of adding a humanlike avatar, rejecting H5(a) and H5(b), we did not find significant differences between the humanlike-AI and the plain-AI conditions on participants' certainty or social presence. With a small- to medium-size effect that was close to significance, we found the humanlike AI system induced more social presence ($M_{\text{difference}} = 0.60$, $p = .099$, 95% CI [−0.11 to 1.31], 90% CI [0.003 to 1.20], Cohen's $d = 0.359$) than the plain AI system. In answering RQ2, we found the humanlike avatar decreased participants' articulation rate relative to the plain AI system ($M_{\text{difference}}$ = −0.23, p = .005, 95% CI [−0.39 to −0.07], Cohen's $d = -0.627$). No significant differences were observed on other speech fluency measures.

To test H2 and H4 regarding the mechanisms underlying the effects of AI evaluation (vs. human evaluation) on speech fluency, we performed a series of mediation analyses with Model 4 of PROCESS (v4.1, Hayes, 2022), with certainty (H2) and social presence (H4) as parallel mediators, and each speech fluency measure as a dependent variable while controlling for the same covariates. We requested 5,000 samples of bootstrapping with both 95% CI and 90% CI.

Table 3 summarizes the mediation results. Rejecting H2 and H4, we did not find any significant indirect effects of AI

Table 1. Definitions and operationalization of speech fluency

Variables	Definitions	Operationalization
Speech rate	Amount of speech produced per unit of time allocated to a speaking task	The total number of syllables divided by the total time allocated to answering the six interview questions, in seconds
Articulation rate	Amount of speech produced per unit of phonation time	The total number of syllables divided by phonation time (i.e., the time from the first phoneme to the last syllable excluding silent pause intervals, in seconds) answering the six interview questions
Onset latency	The time used to construct an answer before speaking	The average time duration from the beginning of the recording to the first phoneme spoken in one's answers to the six interview questions, in seconds
Avg. frequency of silent pauses	Number of silent pauses per unit of time	The total number of silent pauses (i.e., no phonation longer than 250 ms, De Jong & Bosker, 2013; Goldman-Eisler, 1968; Tavakoli & Wright, 2020) divided by the total phonation time answering the six interview questions
Avg. frequency of filled pauses	Number of filled pauses per unit of time	The total number of filled pauses (e.g., "er" and "um") divided by the total phonation time answering the six interview questions

Table 2. Analysis of covariance results

	F statistics			Experimental conditions		
	$F(2, 126)$	p	η^2	Human	AI	Humanlike AI
Certainty	3.12	.047	0.047	4.17 (0.18)^b	3.56 (0.17)^a	3.80 (0.18)
Social presence	2.98	.055	0.045	4.24 (0.26)^b	3.40 (0.24)^a	4.00 (0.26)^c
Speed-related measures of speech fluency						
Speech rate	0.77	.47	0.012	3.22 (0.07)	3.33 (0.07)	3.23 (0.07)
Articulation rate	4.29	.016	0.064	4.45 (0.06)^c	4.61 (0.05)^a	4.38 (0.06)^b
Breakdown-related measures of speech fluency						
Onset latency	0.44	.64	0.007	1.52 (0.09)	1.46 (0.08)	1.41 (0.09)
Avg. frequency of silent pauses	1.41	.25	0.022	0.49 (0.03)	0.50 (0.02)	0.44 (0.03)
Avg. frequency of filled pauses	0.46	.63	0.007	0.14 (0.01)	0.15 (0.01)	0.15 (0.01)

Notes:
The format of mean comparisons is mean (standard error). Within rows, means with different superscripts [a, b] differ at $p < .05$; means with different superscripts [a, c] differ at $p < .10$.

Table 3. Parallel mediation results

Speech disfluency	Mediator	Indirect effects (reference group = plain AI system)					
		Human			Humanlike AI		
		B(Boot*SE*)	95% CI	90% CI	*B*(Boot*SE*)	95% CI	90% CI
Speech rate	Certainty	−0.02 (0.03)	[−0.08 to 0.03]	[−0.08 to 0.02]	−0.01(0.01)	[−0.04 to 0.01]	[−0.03 to 0.01]
	Social presence	−0.04 (0.03)	[−0.10 to 0.004]	**[−0.09 to −0.001]**	−0.03(0.02)	[−0.07 to 0.005]	[−0.06 to 0.002]
Articulation rate	Certainty	−0.03 (0.02)	[−0.08 to 0.01]	[−0.07 to 0.001]	−0.01(0.01)	[−0.04 to 0.01]	[−0.04 to 0.001]
	Social presence	−0.001 (0.02)	[−0.04 to 0.04]	[−0.03 to 0.03]	−0.001(0.01)	[−0.03 to 0.03]	[−0.02 to 0.02]
Onset latency	Certainty	−0.04 (0.03)	[−0.12 to 0.02]	[−0.10 to 0.01]	−0.01(0.02)	[−0.06 to 0.01]	[−0.05 to 0.01]
	Social presence	0.03 (0.03)	[−0.02 to 0.11]	[−0.01 to 0.09]	0.02(0.02)	[−0.02 to 0.08]	[−0.01 to 0.06]
Avg. frequency of silent pauses	Certainty	−0.01 (0.01)	[−0.02 to 0.01]	[−0.02 to 0.01]	−0.003(0.005)	[−0.01 to 0.01]	[−0.01 to 0.004]
	Social presence	0.01 (0.01)	[−0.002 to 0.04]	**[0.0001 to 0.03]**	0.01(0.01)	[−0.002 to 0.03]	[−0.001 to 0.02]
Avg. frequency of filled pauses	Certainty	−0.003 (0.005)	[−0.01 to 0.01]	[−0.01 to 0.01]	−0.001(0.002)	[−0.01 to 0.003]	[−0.01 to 0.002]
	Social presence	0.001 (0.005)	[−0.01 to 0.01]	[−0.01 to 0.01]	0.001(0.003)	[−0.01 to 0.01]	[−0.004 to 0.01]

Bold represents the only values that have passed the statistical threshold.

evaluation (vs. human evaluation) on speech fluency via certainty or social presence. We found two small, close-to-significance indirect effects of AI evaluation via social presence such that by reducing social presence, AI evaluation increased speech rate ($B = 0.04$, Boot$SE = 0.03$, 95% CI [−0.004 to 0.10], 90% CI [0.001 to 0.09]) and decreased the average frequency of silent pauses ($B = -0.01$, Boot$SE = 0.01$, 95% CI [−0.04 to 0.002], 90% CI [−0.03 to −0.0001]).

Summary of findings

In sum, compared with human evaluation, AI evaluation induced higher uncertainty (H1 supported) and lower social presence (H3 supported) with significant, medium-size effects and increased articulation rate with a small- to medium-size effect (nonsignificant at $\alpha = 0.05$ level, significant at $\alpha = 0.10$ level), but did not influence other fluency measures (RQ1, with nonsignificant, small-size effects). Uncertainty and social presence did not significantly mediate the effects of AI evaluation on any of the fluency measures (H2 and H4 not supported). However, with nonsignificant, small effects, AI evaluation increased speech rate and decreased the frequency of silent pauses via social presence (significant at $\alpha = 0.10$ level).

Compared with the plain AI system, adding a humanlike avatar did not significantly reduce uncertainty (H5a not supported), increased social presence with a nonsignificant, small- to medium-size effect (H5b not supported; significant at $\alpha = 0.10$ level), and decreased articulation rate with a significant, medium-size effect (RQ2), but did not affect other measures of speech fluency.

Discussion

Replacing humans with machines in communication changes the communication context and may influence individuals' psychological experience and message production (Gambino & Liu, 2022). With AI technology widely adopted in evaluating job candidates and informing hiring decisions, the current study serves as an initial effort that experimentally tested the effects of AI evaluation in job interviews on the perceptions and performance of job candidates. By shedding light on the underlying mechanisms of speech production in HMC, the findings contribute to the theoretical understanding of HMC and also have practical implications for AI system design and professional training.

Theoretical implications

First, we examined individuals' psychological experience of being evaluated by an AI system relative to human recruiters in terms of uncertainty and social presence during mock job interviews. We found participants perceive greater uncertainty about AI evaluation than human evaluation with a medium-size difference, which corroborates the findings from existing research using qualitative (e.g., Kim & Heo, 2021; Mirowska & Mesnet, 2022) and vignette-based methods (e.g., Schick & Fischer, 2021). This finding is also consistent with HMC research examining other types of machine agents, such as social robots (Edwards et al., 2016), which in general found individuals have higher uncertainty about machine agency than human agency, potentially because individuals could use anthropocentric knowledge to reduce their uncertainty about other humans, but the knowledge is less applicable and accessible when interacting with machines (Epley et al., 2007). We predicted that adding humanlike avatars would render anthropocentric knowledge more accessible so as to reduce individuals' uncertainty about AI, but the results did not support this prediction, which is inconsistent with previous research that found anthropomorphic cues reduce uncertainty (Ososky et al., 2013). We speculate that such an inconsistency could be due to contextual factors. According to the three-factor theory of anthropomorphism (Epley et al., 2007), anthropomorphism happens when people are motivated to reduce uncertainty. During job interviews, however, uncertainty reduction might not be individuals' primary goal, and therefore, they might not be sufficiently motivated to anthropomorphize the AI system to reduce uncertainty.

We found participants in the plain-AI condition reported lower social presence than those in the human-evaluation condition with a medium-size difference, which replicated previous research findings that individuals experience lower social presence in AI-based interviews (Langer et al., 2019), and lower social presence with machine agency than with human agency in general (e.g., Edwards et al., 2021; Fox et al., 2015; Liu, 2021). Adding humanlike avatars to the plain AI system increased social presence with a small- to medium-size effect (significant at $\alpha = 0.10$). These results support Blascovich's (2002) Threshold Model of Social Influence that says humans evoke greater social presence and social responses than artificial entities; the realism of the artificial entities in terms of forms or behaviors can increase social presence to some extent.

Second, in addition to examining individuals' experience of AI-based interviews, one major innovation of the current study is the examination of communication-related features, which was difficult to test in vignette-based or retrospection-based research (Langer & Landers, 2021). We proposed and tested two theory-driven mechanisms that could potentially influence speech fluency in different directions. The first mechanism, as informed by theories and research in message production under uncertainty, hypothesized that AI, by inducing higher uncertainty, could debilitate speech production. Mediation analyses, however, did not reveal any significant indirect effects of AI evaluation via uncertainty on speech fluency measures, which suggests that uncertainty about AI evaluation might not be a dominant process during individuals' speech production in the context of job interviews.

The finding that uncertainty about AI not affecting speech fluency has a methodological implication for HMC research. In previous research that found uncertainty, and relatedly, the opaqueness of AI evaluation as a severe concern, participants either retrospectively reflected on or imagined the experience of AI-based interviews (e.g., Kim & Heo, 2021; Mirowska & Mesnet, 2022; Schick & Fischer, 2021). When asked to recall or imagine, individuals have the mental capacity to attend to and evaluate the features of AI-based interviews. However, during job interviews, individuals may become preoccupied with processing and responding to the interview questions and not motivated or capable to attend to all features of AI (e.g., its opaqueness). Hence, the methodology used to study HMC matters; participants' experience that is salient in non-interactive studies may become less relevant during real-time interaction due to the different attentional processes.

In addition to uncertainty, we also examined another mechanism—reduced social presence—that could potentially facilitate message production in AI-based interviews. Mediation analyses found AI evaluation, by reducing social presence, increased speech rate and reduced silent pauses with small-size effects (significant at $\alpha = 0.10$). In addition, AI evaluation (vs. human evaluation) induced a higher articulation rate with a small- to medium-size effect (significant at $\alpha = 0.10$). Adding a humanlike avatar to the plain interface significantly decreased articulation rate with a medium-size effect. This suggests both human and humanlike AI evoked social facilitation responses, such that when individuals perceived another human or humanlike entity observing them, their task performance (i.e., speech production) was undermined (Zajonc, 1965), which again lends support to Blascovich's (2002) Threshold Model of Social Influence. Because articulation rate is negatively associated with cognitive load (Tavakoli & Skehan, 2005), these findings seem to suggest that individuals have lower, rather than higher difficulty, in conversing with machines than with humans, which is consistent with previous findings that individuals felt easier to disclose their health issues to chatbots than to humans (Bae Brandtzæg et al., 2021; Lucas et al., 2014).

Nevertheless, we would like to point out that although AI evaluation increased the articulation rate, this effect was not mediated by social presence. As previously mentioned, articulation rate is a pure measure of speed whereas speech rate is a composite measure of both speed and breakdown (Tavakoli & Wright, 2020). Based on our results, it seems that social presence only affects breakdown fluency (as indicated by its effects on speech rate and silent pauses), which is linked to the conceptualization stage of speech production when pre-verbal plans are formed (Levelt, 1999), such as deciding what to say next (i.e., discourse plan, Rose, 2017), but does not affect articulation. The finding that human evaluation induces greater social presence and thereby interrupts planning lends support to the goal-complexity perspective forwarded by Gambino and Liu (2022), which argues that communication with humans could be more cognitively demanding than with machines for the activation of additional secondary goals that are tied to unique human capacities, such as concerns for another person's emotional responses, which does not apply to machines. As goal structure becomes more complex, so does planning, which may then interfere with speech production (Knowlton & Berger, 1997).

Taken together, one major implication of our findings for HMC research is that a global contrast between HMC vs. human communication might not be sufficient in understanding the nuances in HMC. Contemporary research often approaches HMC using human communication as a benchmark (Guzman & Lewis, 2020). Yet, as shown in the present findings, it is oversimplified to conclude that HMC is similar or different from human communication, as we found some aspects of HMC resembled human communication (i.e., the social facilitation effects), while others did not (i.e., effects of uncertainty on message production). In the current context of job interviews, individuals' impression management goal might be more salient than the uncertainty reduction goal, which may explain why presence-related features and processes had a larger impact on speech production. Hence, we join some scholars in arguing that in HMC, the salience and effects of certain psychological processes depend on the characteristics of the technology, the individuals (Gambino et al., 2020), and the interaction context (Gambino & Liu, 2022). As a remedy, we recommend a richer analysis of context and applying a variable-based approach (Nass & Mason, 1990) in identifying and testing characteristics of machines and individuals that are relevant to the context, which might help develop new theoretical understandings of HMC.

Practical implications

As AI evaluation heightened uncertainty and lowered social presence, such psychological experiences may be reflected in job interviewees' verbal and nonverbal behaviors. For instance, a suitable candidate who would have performed well in a human-led interview may show signs of uncertainty and a lack of engagement (due to low social presence) when responding to AI and be misjudged as incompetent or disinterested. Hence, rather than applying human–human communication metrics uncritically, AI-based systems should be designed to be context-aware and evaluate candidates' performance with context-specific metrics.

Relatedly, another problem with insufficient consideration of the contextual influence is discrimination against underrepresented populations. The uncritical application of evaluation criteria encoded and normalized by AI systems developed based on privileged individuals (e.g., those who are already hired) may reinforce the existing systematic inequality (Whittaker et al., 2019). For example, using fluency-related metrics as evaluation criteria reflects an ableist perspective (Fruchterman & Mellea, 2018). While calling for building context-aware AI systems, we also call for considering the idiosyncratic responses that individuals have to AI-based interviews to promote equity and inclusion in the workplace.

Given the extant practice of AI-based interviews, we also have suggestions for job seekers and institutions that offer training on interview preparation. Based on the finding that AI evaluation increased uncertainty and decreased social presence, we recommend using strategies to overcome uncertainty (e.g., by learning more about AI and practicing more) and to compensate for the low social presence during AI-based interviews (e.g., by acting more engaged than one would normally do).

Limitations and future directions

The current study has a few limitations to consider. First, except for increasing articulation rate (significant at $\alpha = 0.10$), we did not find effects of AI evaluation on other measures of speech fluency. This might be due to participants' low involvement, as the stakes in the current task (i.e., having a mock interview, as opposed to a real job interview) were low (Langer et al., 2019). Also, with a sample size of 134, this study is underpowered to detect small effects. Future research should replicate and extend the current study with a larger sample size with a more realistic setup.

Second, we told participants that they would be evaluated either by an AI system or by a human and only examined the effects of machine agency in evaluating, rather than conducting the interview (e.g., generating interview questions). In reality, the level of automation and the role of humans can vary on a spectrum. For example, humans can still set parameters for the AI system or make the final decisions in light of AI's recommendations (e.g., Gonzalez et al., 2022; Köchling et al., 2022). Future research can explore the effects of machine agency of various quantities and qualities to better assess the trade-offs involved in using AI for hiring.

Third, we only used two avatars of particular demographics, which limits the generalizability of the effects of humanlike avatars observed in the current study. Future research should examine whether avatars' characteristics (e.g., gender and ethnicity) influence individuals' perceptions and performance during job interviews.

Fourth, we measured participants' self-reported knowledge of AI to account for the effects of one's familiarity with AI on their performance. Future research may approach this more directly by measuring their previous experience with AI-based agents (e.g., Siri and Alexa).

We would also like to caution that the meanings of speech metrics may be nuanced. For example, greater fluency in terms of speaking fast and having fewer silent pauses could be symptoms of anxiety (e.g., Siegman, 1978) or failing to consider the listener. Moreover, speech fluency is not the only dimension of interview performance; communication difficulty may also be manifested in the content of participants' answers. In this study, however, we instructed participants to think about their dream jobs during the mock interviews and hence might have introduced excessive variance into the content of their answers. Future research could explore whether and how AI evaluation influences the content of interviewees' responses and other aspects of communication (e.g., linguistic styles and non-verbal cues).

Conclusion

In this study, we found AI evaluation (vs. human evaluation) in job interviews led interviewees to perceive higher uncertainty and lower social presence and to have a higher articulation rate (significant at $\alpha = 0.10$). By reducing social presence, AI evaluation increased speech rate and reduced frequency of silent pauses (both indirect effects significant at $\alpha = 0.10$). Findings shed light on individuals' psychological experience and speech production in HMC in the context of hiring, based on which, we suggest employers and AI designers calibrate and account for AI-induced effects and recommend job seekers reduce uncertainty about AI and compensate for the reduced social presence during AI-based interviews.

Supplementary material

Supplementary material is available at *Journal of Computer-Mediated Communication.*

Data availability

The data underlying this article are available upon reasonable request.

Funding

Data collection was funded by the first and third authors' faculty development fund at California State University, Los Angeles.

Conflicts of interest: The authors declare that there is no conflict of interest.

Acknowledgments

We gratefully acknowledge the insights from Dr James P. Dillard and Dr Denise H. Solomon and the constructive feedback from the editor and reviewers.

Open science framework badges

Open Materials

The components of the research methodology needed to reproduce the reported procedure and analysis are publicly available for this article.

References

Acikgoz, Y., Davison, K. H., Compagnone, M., & Laske, M. (2020). Justice perceptions of artificial intelligence in selection. *International Journal of Selection and Assessment*, *28*(4), 399–416. https://doi.org/10.1111/ijsa.12306

Adepu, Y., Boga, V. R., & Sairam, U. (2020). Interviewee performance analyzer using facial emotion recognition and speech fluency recognition. In A. Patil (Ed.), *2020 IEEE international conference for innovation in technology* (pp. 1–5). IEEE. https://doi.org/10.1109/INOCON50539.2020.9298427

Araujo, T., Helberger, N., Kruikemeier, S., & De Vreese, C. H. (2020). In AI we trust? Perceptions about automated decision-making by artificial intelligence. *AI & Society*, *35*(3), 611–623. https://doi.org/10.1007/s00146-019-00931-w

Bae Brandtzæg, P. B., Skjuve, M., Kristoffer Dysthe, K. K., & Følstad, A. (2021). When the social becomes non-human: Young people's perception of social support in Chatbots. In P. Bjørn & S. Drucker (Eds.), *Proceedings of the 2021 CHI conference on human factors in computing systems* (pp. 1–13). Association for Computing Machinery. https://doi.org/10.1145/3411764.3445318

Berger, C. R. (1988). Planning, affect, and social action generation. In L. Donohew, H. E. Sypher, & E. T. Higgins (Eds.), *Communication, social cognition, and affect* (pp. 93–116). Lawrence Erlbaum Associates.

Berger, C. R. (1995). A plan-based approach to strategic communication. In D. E. Hewes (Ed.), *The cognitive bases of interpersonal communication* (pp. 141–179). Erlbaum.

Berger, C. R. (1997). Producing messages under uncertainty. In J. O. Greene (Ed.), *Message production: Advances in communication theory* (pp. 221–244). Lawrence Erlbaum.

Berger, C. R., & Bradac, J. J. (1982). *Language and social knowledge: Uncertainty in interpersonal relations*. Edward Arnold.

Berger, C. R., & Jordan, J. M. (1992). Planning sources, planning difficulty and verbal fluency. *Communications Monographs*, *59*(2), 130–149. https://doi.org/10.1080/03637759209376257

Blascovich, J. (2002). A theoretical model of social influence for increasing the utility of collaborative virtual environments. In W. Broll, C. Greenhalgh, & E. Churchill (Eds.), *Proceedings of the 4th international conference on collaborative virtual environments* (pp. 25–30). Association for Computing Machinery. https://doi.org/10.1145/571878.571883

Breil, S. M., Osterholz, S., Nestler, S., & Back, M. D. (2021) Contributions of nonverbal cues to the accurate judgment of personality traits. In T. D. Letzring, and J. S. Spain (Eds.), *The Oxford handbook of accurate personality judgment* (pp. 195–218). Oxford University Press. https://doi.org/10.1093/oxfordhb/9780190912529.013.13

Brennan, S. E., & Williams, M. (1995). The feeling of another's knowing: Prosody and filled pauses as cues to listeners about the metacognitive states of speakers. *Journal of Memory and Language*, *34*(3), 383–398. https://doi.org/10.1006/jmla.1995.1017

Brosy, J., Bangerter, A., & Mayor, E. (2016). Disfluent responses to job interview questions and what they entail. *Discourse Processes*, *53*(5–6), 371–391. https://doi.org/10.1080/0163853X.2016.1150769

Clatterbuck, G. W. (1979). Attributional confidence and uncertainty in initial interaction. *Human Communication Research*, *5*(2), 147–157. https://doi.org/10.1111/j.1468-2958.1979.tb00630.x

CloudResearch. (2022). *CloudResearch | Online research & participant recruitment made easy*. https://www.cloudresearch.com/

Constantin, K. L., Powell, D. M., & McCarthy, J. M. (2021). Expanding conceptual understanding of interview anxiety and performance: Integrating cognitive, behavioral, and physiological features. *International Journal of Selection and Assessment*, *29*(2), 234–252. https://doi.org/10.1111/ijsa.12326

De Jong, N. H., & Bosker, H. R. (2013). Choosing a threshold for silent pauses to measure second language fluency. In R. Eklund (Ed.), *Proceedings of the 6th workshop on disfluency in spontaneous speech* (pp. 17–20). Royal Institute of Technology.

De Jong, N. H., Pacilly, J., & Heeren, W. (2021). PRAAT scripts to measure speed fluency and breakdown fluency in speech automatically. *Assessment in Education: Principles, Policy & Practice*, *28*(4), 456–476, https://doi.org/10.1080/0969594X.2021.1951162

Dillard, J. P., Segrin, C., & Harden, J. M. (1989). Primary and secondary goals in the production of interpersonal influence messages. *Communications Monographs*, *56*(1), 19–38. https://doi.org/10.1080/03637758909390247

Edwards, C., Edwards, A., Albrehi, F., & Spence, P. (2021). Interpersonal impressions of a social robot versus human in the context of performance evaluations. *Communication Education*, *70*(2), 165–182. https://doi.org/10.1080/03634523.2020.1802495

Edwards, C., Edwards, A., Spence, P. R., & Westerman, D. (2016). Initial interaction expectations with robots: Testing the human-to-human interaction script. *Communication Studies*, *67*(2), 227–238. https://doi.org/10.1080/10510974.2015.1121899

Edwards, C., Stoll, B., Edwards, A., Spence, P., & Gambino, A. (2018). "I'll present to the human": Effects of a robot evaluator on anticipatory public speaking anxiety. In A. Guzman (Eds.), *Human–machine communication: Rethinking communication, technology, and ourselves*. Peter Lang.

Epley, N., Waytz, A., & Cacioppo, J. T. (2007). On seeing human: A three-factor theory of anthropomorphism. *Psychological Review*, *114*(4), 864–886. https://doi.org/10.1037/0033-295X.114.4.864

Faul, F., Erdfelder, E., Buchner, A., & Lang, A. G. (2009). Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods*, *41*(4), 1149–1160. https://doi.org/10.3758/BRM.41.4.1149

Fox, J., Ahn, S. J., Janssen, J. H., Yeykelis, L., Segovia, K. Y., & Bailenson, J. N. (2015). Avatars versus agents: A meta-analysis quantifying the effect of agency on social influence. *Human-Computer Interaction*, *30*(5), 401–432. https://doi.org/10.1080/07370024.2014.921494

Fruchterman, J., & Mellea, J. (2018). Expanding employment success for people with disabilities. *Benetech*. https://benetech.org/about/resources/expanding-employment-success-for-people-with-disabilities-2/

Gambino, A., Fox, J., & Ratan, R. A. (2020). Building a stronger CASA: Extending the computers are social actors paradigm. *Human–Machine Communication*, *1*, 71–86. https://doi.org/10.30658/hmc.1.5

Gambino, A., & Liu, B. (2022). Considering the context to build theory in HCI, HRI, and HMC: Explicating differences in processes of communication and socialization with social technologies. *Human–Machine Communication*, *4*, 111–130. https://doi.org/10.30658/hmc.4.6

Goldman-Eisler, F. (1968). *Psycholinguistics: Experiments in spontaneous speech*. Academic Press.

Gonzalez, M. F., Liu, W., Shirase, L., Tomczak, D. L., Lobbe, C. E., Justenhoven, R., & Martin, N. R. (2022). Allying with AI? Reactions toward human-based, AI/ML-based, and augmented hiring processes. *Computers in Human Behavior*, *130*, 107179. https://doi.org/10.1016/j.chb.2022.107179

Greene, J. O., & Lindsey, A. E. (1989). Encoding processes in the production of multiple-goal messages. *Human Communication Research*, *16*(1), 120–140. https://doi.org/10.1111/j.1468-2958.1989.tb00207.x

Guzman, A. L., & Lewis, S. C. (2020). Artificial intelligence and communication: A human–machine communication research agenda. *New Media & Society*, *22*(1), 70–86. https://doi.org/10.1177/1461444819858691

Harwell, D (2019, November 6). A face-scanning algorithm increasingly decides whether you deserve the job. *The Washington Post*. https://www.washingtonpost.com/technology/2019/10/22/ai-hiring-face-scanning-algorithm-increasingly-decides-whether-you-deserve-job/

Hayes, A. F. (2022). *PROCESS* (Version 4.1) [Computer software]. https://www.processmacro.org/download.html

Ho, C. C., & MacDorman, K. F. (2010). Revisiting the uncanny valley theory: Developing and validating an alternative to the Godspeed indices. *Computers in Human Behavior*, *26*(6), 1508–1518. https://doi.org/10.1016/j.chb.2010.05.015

Hoyt, C. L., Blascovich, J., & Swinth, K. R. (2003). Social inhibition in immersive virtual environments. *Presence: Teleoperators & Virtual Environments*, *12*(2), 183–195. https://doi.org/10.1162/105474603321640932

Jaser, Z., Petrakaki, D., Starr, R., & Oyarbide-Magaña, E. (2022). Where automated job interviews fall short. *Harvard Business Review*. https://hbr.org/2022/01/where-automated-job-interviews-fall-short

Johnson-Laird, P. N. (1980). Mental models in cognitive science. *Cognitive Science*, *4*(1), 71–115. https://doi.org/10.1016/S0364-0213(81)80005-5

Kiesler, S., & Goetz, J. (2002). Mental models of robotic assistants. In L. Terveen & D. Wixon (Eds.), *CHI'02 extended abstracts on human factors in computing systems* (pp. 576–577). Association for Computing Machinery. https://doi.org/10.1145/506443.506491

Kim, J.-Y., & Heo, W. (2021). Artificial intelligence video interviewing for employment: Perspectives from applicants, companies, developer and academicians. *Information Technology & People*, *35*(3), 861–878. https://doi.org/10.1108/ITP-04-2019-0173

Knobloch, L. K., & Solomon, D. H. (2005). Relational uncertainty and relational information processing: Questions without answers? *Communication Research*, *32*(3), 349–388. https://doi.org/10.1177/0093650205275384

Knobloch, L. K. (2006). Relational uncertainty and message production within courtship: Features of date request messages. *Human Communication Research*, *32*(3), 244–273. https://doi.org/10.1111/j.1468-2958.2006.00275.x

Knowlton, S. W., & Berger, C. R. (1997). Message planning, communication failure, and cognitive load: Further explorations of the hierarchy principle. *Human Communication Research*, *24*(1), 4–30. https://doi.org/10.1111/j.1468-2958.1997.tb00585.x

Köchling, A., Wehner, M. C., & Warkocz, J. (2022). Can I show my skills? Affective responses to Artificial Intelligence in the recruitment process. *Review of Managerial Science*. https://doi.org/10.1007/s11846-021-00514-4

Langer, M., König, C. J., & Krause, K. (2017). Examining digital interviews for personnel selection: Applicant reactions and interviewer ratings. *International Journal of Selection and Assessment*, *25*(4), 371–382. https://doi.org/10.1111/ijsa.12191

Langer, M., König, C. J., & Papathanasiou, M. (2019). Highly automated job interviews: Acceptance under the influence of stakes. *International Journal of Selection and Assessment*, *27*(3), 217–234. https://doi.org/10.1111/ijsa.12246

Langer, M., König, C. J., & Hemsing, V. (2020a). Is anybody listening? The impact of automatically evaluated job interviews on impression management and applicant reactions. *Journal of Managerial Psychology*, *35*(4), 271–284. https://doi.org/10.1108/JMP-03-2019-0156

Langer, M., König, C. J., Sanchez, D. R.-P., & Samadi, S. (2020b). Highly automated interviews: Applicant reactions and the organizational context. *Journal of Managerial Psychology*, *35*(4), 301–314. https://doi.org/10.1108/jmp-09-2018-0402

Langer, M., & Landers, R. N. (2021). The future of artificial intelligence at work: A review on effects of decision automation and augmentation on workers targeted by algorithms and third-party observers. *Computers in Human Behavior*, *123*, 106878. https://doi.org/10.1016/j.chb.2021.106878

Lee, K. M., Peng, W., Jin, S. A., & Yan, C. (2006). Can robots manifest personality?: An empirical test of personality recognition, social responses, and social presence in human–robot interaction. *Journal of Communication*, *56*(4), 754–772. https://doi.org/10.1111/j.1460-2466.2006.00318.x

Levelt, W. (1999). Language production: A blueprint of the speaker. In C. Brownand P. Hagoort (Eds.), *Neurocognition of language* (pp. 83–122). Oxford University Press.

Liu, B. (2021). In AI we trust? Effects of agency locus and transparency on uncertainty reduction in human–AI interaction. *Journal of Computer-Mediated Communication*, *26*(6), 384–402. https://doi.org/10.1093/jcmc/zmab013

Longoni, C., Bonezzi, A., & Morewedge, C. K. (2019). Resistance to medical artificial intelligence. *Journal of Consumer Research*, *46*(4), 629–650. https://doi.org/10.1093/jcr/ucz013

Lucas, G. M., Gratch, J., King, A., & Morency, L. P. (2014). It's only a computer: Virtual humans increase willingness to disclose. *Computers in Human Behavior*, *37*, 94–100. https://doi.org/10.1016/j.chb.2014.04.043

Luger, G. & Stubblefield, R. (1997). *Artificial intelligence: Structures and strategies for complex problem solving*. Addison-Wesley.

Martín-Raugh, M. P., Leong, C. W., Roohr, K. C., & Chen, X. (2023). Perceived conscientiousness and openness to experience mediate the relationship between vocal and visual features and hiring decision in the interview. *Computers in Human Behavior Reports*, *10*, 100272. https://doi.org/10.1016/j.chbr.2023.100272

McCarthy, J., & Goffin, R. (2004). Measuring job interview anxiety: Beyond weak knees and sweaty palms. *Personnel Psychology*, *57*(3), 607–637. https://doi.org/10.1111/j.1744-6570.2004.00002.x

Miller, G. R., & Steinberg, M. (1975). *Between people: A new analysis of interpersonal communication*. Science Research Associates.

Mirowska, A., & Mesnet, L. (2022). Preferring the devil you know: Potential applicant reactions to Artificial Intelligence evaluation of interviews. *Human Resource Management Journal*, *32*(2), 364–383. https://doi.org/10.1111/1748-8583.12393

Mittelstadt, B. D., Allo, P., Taddeo, M., Wachter, S., & Floridi, L. (2016). The ethics of algorithms: Mapping the debate. *Big Data & Society*, *3*(2), 1–21. https://doi.org/10.1177/2053951716679679

Naim, I., Tanveer, Md. I., Gildea, D., & Hoque, M. E. (2018). Automated analysis and prediction of job interview performance. *IEEE Transactions on Affective Computing*, *9*(2), 191–204. https://doi.org/10.1109/TAFFC.2016.2614299

Nass, C., & Mason, L. (1990). On the study of technology and task: A variable-based approach. In J. Fulk & C. Steinfield (Eds.), *Organizations and communication technology* (pp. 46–67). Sage.

Oh, C. S., Bailenson, J. N., & Welch, G. F. (2018). A systematic review of social presence: Definition, antecedents, and implications. *Frontiers in Robotics and AI*, *5*. https://doi.org/10.3389/frobt.2018.00114

Ososky, S., Philips, E., Schuster, D., & Jentsch, F. (2013). A picture is worth a thousand mental models: Evaluating *human understanding of robot teammates*. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *57*(1), 1298–1302. https://doi.org/10.1177/1541931213571287

Pickard, M. D., & Roster, C. A. (2020). Using computer automated systems to conduct personal interviews: Does the mere presence of a human face inhibit disclosure? *Computers in Human Behavior*, *105*, 106197. https://doi.org/10.1016/j.chb.2019.106197

Robles, M. M. (2012). Executive perceptions of the top 10 soft skills needed in today's workplace. *Business and Professional Communication Quarterly*, *75*(4), 453–465. https://doi.org/10.1177/108056991246040

Rose, R. (2017). Silent and filled pauses and speech planning in first and second language production. In R. Eklund & R. Rose (Eds.), *Proceedings of disfluency in spontaneous speech 2017* (pp. 49–52). Royal Institute of Technology.

Rouse, W. B., & Morris, N. M. (1986). On looking into the black box: Prospects and limits in the search for mental models. *Psychological Bulletin*, *100*(3), 349–363. http://dx.doi.org/10.1037/0033-2909.100.3.349

Schick, J., & Fischer, S. (2021). Dear computer on my desk, which candidate fits best? An assessment of candidates' perception of assessment quality when using AI in personnel selection. *Frontiers in Psychology*, *12*(739711). https://doi.org/10.3389/fpsyg.2021.739711

Siegman, A. W. (1978). The meaning of silent pauses in the initial interview. *The Journal of Nervous and Mental Disease*, *166*(9), 642–654. https://doi.org/10.1097/00005053-197809000-00004

Şimşek, Ö. F., & Yalınçetin, B. (2010). I feel unique, therefore I am: The development and preliminary validation of the personal sense of uniqueness (PSU) scale. *Personality and Individual Differences*, *49*(6), 576–581. https://doi.org/10.1016/j.paid.2010.05.006

Suen, H. Y., Chen, M. Y. C., & Lu, S. H. (2019). Does the use of synchrony and artificial intelligence in video interviews affect interview ratings and applicant attitudes? *Computers in Human Behavior*, *98*, 93–101. https://doi.org/10.1016/j.chb.2019.04.012

Synthesia. (2022). *Synthesia | #1 AI video generation platform*. https://www.synthesia.io/

Tavakoli, P. & Skehan, P. (2005). Strategic planning, task structure, and performance testing. In R. Ellis (Ed.), *Planning and task performance* (pp. 239–273). John Benjamins.

Tavakoli, P., & Wright, C. (2020). *Second language speech fluency: From research to practice*. Cambridge University Press.

Whittaker, M., Alper, M., Bennett, C. L., Hendren, S., Kaziunas, L., Mills, M., Morris, M. R., Rankin, J., Rogers, E., Salas, M., & West, S. M. (2019). *Disability, bias, and AI*. AI Now Institute. https://ainowinstitute.org/disabilitybiasai-2019.pdf

Zajonc, R. B. (1965). Social facilitation: A solution is suggested for an old unresolved social psychological problem. *Science*, *149*(3681), 269–274. https://doi.org/10.1126/science.149.3681.269