The Effects of AI Interviewers' Gender and Professionalism on Applicant Reactions: The Mediating Role of Trust*

Wang Lina, Zheng Lu, Alexander Scott English, Zhao Teng, Li Hairong, Jiang Yan

2026-08-24

ChinaRxiv

Abstract

AI interviews are becoming an important tool in organizational recruitment processes, and an increasing number of job applicants need to interact with AI interviewers rather than human interviewers. Against this background, understanding how the external image characteristics of AI interviewers influence applicant reactions is particularly critical. Based on similarity-attraction theory, this paper, through two interview simulation experiments and an experience sampling study based on real job applicants, finds that: (1) gender congruence between AI interviewers and applicants can enhance applicants' trust in the interviewer and further reduce their perceptions of process creepiness and privacy concerns; (2) interviewer image professionalism plays a moderating role. When interviewer image professionalism is low, applicants rely more on gender congruence cues, and the indirect effects of gender congruence on process creepiness and privacy concerns through interviewer trust are more pronounced. This study extends the application of similarity-attraction theory to human-computer interaction and helps deepen understanding of the influence of AI interviewers' external image on applicants' emotional and cognitive reactions, providing implications for the design of AI interviews.

Full Text

The Effects of AI Interviewers' Gender and Professionalism on Applicant Reactions: The Mediating Role of Trust*

Wang Lina[1] Zheng Lu[1] Alexander Scott English[2,3] Zhao Teng[4] Li Hairong[5] Jiang Yan[1]

([1] School of Management, Huazhong University of Science and Technology, Wuhan 430074)

([2] Wenzhou-Kean University Center for Chinese-American Cultural and Economic Studies, [3] Center for Cultural Behavior Research, College of Liberal Arts, Wenzhou-Kean University, Wenzhou 325060)

([4] Business School, Nankai University, Tianjin 300071) ([5] School of Labor and Human Resources, Renmin University of China, Beijing 100872)

Abstract Artificial intelligence interviews are becoming an important tool in organizational recruitment processes, and increasing numbers of job applicants need to interact with AI interviewers rather than human interviewers. Against this background, understanding how the external image characteristics of AI interviewers affect applicant reactions is particularly critical. Drawing on similarity-attraction theory, this paper, through two interview-simulation experiments and one experience-sampling study based on real job applicants, finds that: (1) gender congruence between the AI interviewer and the applicant can

enhance applicants' trust in the interviewer and further reduce their process creepiness and privacy concerns; (2) the professionalism of the interviewer' s image plays a moderating role. When the interviewer' s image professionalism is low, applicants rely more on the cue of gender congruence, and the indirect effects of gender congruence on process creepiness and privacy concerns through trust in the interviewer are more pronounced. This research extends the application of similarity-attraction theory to human–machine interaction, helps explain how the external image of AI interviewers influences applicants' emotional and cognitive reactions, and provides a reference for the design of AI interviews.

Keywords AI interviewer, trust, image professionalism, applicant reactions, similarity-attraction theory

Classification number B849

Received: 2025-12-15

* Supported by the General Program of the National Natural Science Foundation of China (72472057), the Fundamental Research Funds for the Central Universities (HUST: 2024WKYXQN026), the Key Program of the National Natural Science Foundation of China (72132001), the Young Scientists Fund of the National Natural Science Foundation of China (72101257, 72401147), and the 2026 Wenzhou Municipal Philosophy and Social Sciences Planning Provincial Think Tank "Chinese-American Cultural and Economic Studies Center" Project (11).

Corresponding author: Zheng Lu, E-mail: zhenglu18@hust.edu.cn

1 Introduction

With the rapid development of Artificial Intelligence (AI) technology, an increasing number of organizations have introduced AI into the interview process (Koch-Bayram & Kaibel, 2024). The Internet Data Center' s *2024 China Recruitment Market Insights White Paper* shows that more than 67% of enterprises have gradually adopted AI interview systems. AI interviews refer to the integration of AI technology with asynchronous video interviews, in which an AI interviewer asks preset structured questions and analyzes job applicants' linguistic features, emotional cues, and behavioral performance to assist in screening candidates (Hunkenschroer & Lütge, 2022). Although AI interviews can improve recruitment efficiency and reduce costs, applicants often exhibit lower trust in them, as well as a stronger sense of process strangeness and privacy concerns (Zhang & Yencha, 2022). On the one hand, when applicants find it difficult to determine what personal information the organization will collect, how the data will be processed, and what type of evaluation method will be used, analysis of voice, facial-expression, and behavioral data can easily trigger privacy concerns (Hunkenschroer & Lütge, 2022). On the other hand, AI interviews lack real-time interaction and feedback with a human being, which

may produce a sense of unfamiliarity, a feeling of being monitored, and an overall uncomfortable process experience (Langer & König, 2018). These negative reactions may further reduce willingness to apply, willingness to accept a job, and organizational attractiveness (McCarthy et al., 2017). Therefore, how to enhance applicants' trust and improve the AI interview experience is an issue that organizations must attend to.

As anthropomorphic AI becomes embedded in recruitment processes, AI interviewers become objects with which applicants directly perceive and interact. When AI presents sufficient social cues, users may apply some interpersonal interaction norms to human–computer interaction (Ahn et al., 2022). In traditional interviews, the interviewer' s external image influences applicant reactions (Uggerslev et al., 2012). Therefore, the external image of an AI interviewer may also become an important cue through which applicants form trust and evaluate the interview process. AI interviewers are virtual AI with a visualized appearance. Unlike embedded AI, which lacks an external appearance, and robotic AI, which has physical hardware, their image is easier to adjust (Glikson & Woolley, 2020). At present, the AI interviewers used by different enterprises differ markedly in appearance, and some systems also support personalized customization of their image (see Appendix). Because virtual AI images are highly malleable, enterprises can iterate and optimize them at relatively low cost. Thus, studying the external image of AI interviewers has important theoretical and practical value.

Gender is an important social-category cue in the external image of AI interviewers. Similar-attraction theory suggests that individuals are more inclined to like and trust objects that have characteristics similar to their own (Byrne, 1997). Interpersonal research has found that racial or gender congruence between recruitment representatives and applicants increases applicants' willingness to accept a job (Young et al., 1997), but interviewers sometimes also prefer opposite-sex applicants (Goldberg, 2005; Graves & Powell, 1995). Similar divergences also exist in human–computer interaction: gender-congruent service robots, medical chatbots, and synthetic voices can improve comfort, communication satisfaction, or credibility (Jin & Eastin, 2024; Lee et al., 2000; Pitardi et al., 2023), whereas opposite-sex robots sometimes receive higher credibility ratings (Siegel et al., 2009). Overall, whether in interpersonal or human–computer interaction, the effect of gender congruence on trust has not yet reached a consistent conclusion.

In recruitment contexts where information is limited and decision time is constrained, individuals rely more on surface-level social cues and tend to choose objects similar to themselves (Abbasi et al., 2024). Because applicants do not understand the operating mechanism and evaluation logic of AI interviewers, gender congruence may

serve as similarity signals, alleviating their concerns about evaluation fairness, professional competence, and algorithmic bias, shortening the psychological distance between humans and AI, and enhancing trust (Mirowska & Arsenyan,

2025). At the same time, the AI interviewer' s face, facial expressions, hairstyle, and clothing can jointly convey professionalism, and the professionalism of the interviewer' s image is also an important basis on which applicants build trust (Wilhelmy et al., 2016), potentially changing the extent to which applicants rely on gender-consistency cues. Therefore, based on similarity-attraction theory (Byrne, 1997), this study examines whether the external image of AI interviewers (gender and professionalism design) affects applicants' trust in the interviewer and further changes their feelings of process creepiness and privacy concerns. The conclusions of this study will help provide actionable design implications for organizations when developing and deploying anthropomorphic recruitment technologies.

1.1 The Effect of AI Interviewer–Applicant Gender Consistency on Trust in the Interviewer

Gender is a fundamental identity dimension that influences social interaction and self-concept (Ridgeway & Smith-Lovin, 1999). Traditional gender stereotypes usually associate men with competence, dominance, and rationality, and women with warmth and empathy (Tay et al., 2014). These stereotypes also extend to AI with gender characteristics (Eagly & Wood, 2012). For example, people generally believe that male AI is more competent, whereas female AI is warmer, and they evaluate AI recommenders that match the product type more positively (Ahn et al., 2022); female chatbots also often receive higher evaluations because they are more humanized and more capable of emotional understanding (Borau, 2021).

Similarity-attraction theory suggests that individuals are more inclined to trust people who have characteristics similar to their own (Byrne, 1997). According to the computers as social actors (CASA) paradigm, when computers or AI present sufficient social cues and are regarded by users as interaction partners, users apply some social rules to human–computer interaction (Gambino et al., 2020; Nass & Moon, 2000). Maeda and Quan-Haase (2024) further interpret human–AI interaction as a quasi-social relationship characterized by asymmetry and unidirectionality: although it lacks the bidirectional emotional reciprocity of traditional interpersonal relationships, users may still perceive AI' s responsiveness and interactivity. Therefore, as AI shifts from a back-end algorithmic tool to an intelligent agent with an external image, natural-language ability, and an interactive role, gender-similarity cues in human–computer interaction may also become an important basis for users to understand and evaluate AI. Research has found that gender-congruent service robots can increase users' comfort during service encounters and enhance their perceived control over the interaction process (Pitardi et al., 2023). Similarly, Jin and Eastin (2024) pointed out that physician chatbots whose gender matches that of users can improve users' communication satisfaction. Lee et al. (2000) further found that gender-consistent computer-synthesized speech significantly enhances its credibility. However, existing research has paid relatively little attention to gender consistency between

interviewers and applicants in AI interview contexts and to its mechanisms of action.

Compared with traditional interpersonal interviews, AI interviews have greater uncertainty, and applicants usually find it difficult to understand the operating mechanisms and evaluation logic of AI interviewers; they are therefore more likely to rely on perceptible external cues when making judgments. Thus, in the high-risk social-evaluation context of an interview, gender consistency helps alleviate applicants' anxiety about evaluation fairness and professional competence, reduce their concerns about algorithmic bias and dehumanized decision-making, and shorten the psychological distance between the two parties, thereby increasing trust in AI interviewers (Mirowska & Arsenyan, 2025). This study therefore proposes the following hypothesis:

Hypothesis 1: Compared with gender inconsistency, gender consistency between the AI interviewer and the applicant will increase applicants' trust in the interviewer.

1.2 The Moderating Role of Interviewer Image Professionalism

Image professionalism is the subjective perception related to professional competence that an individual conveys through external features such as the face (Brownlow, 1992).

When attention is limited or information is insufficient, the professionalism conveyed by external image is more readily processed heuristically than objective professional competence, thereby rapidly influencing trust judgments (Palmer & Peterson, 2016). In recruitment contexts, the professional image of an interviewer can enhance applicants' positive evaluations and favorable attitudes toward the organization (Wilhelmy et al., 2016). Image professionalism has primacy and anchoring effects in the formation of trust; once a professional image constitutes the initial judgment framework, the effects of other cues such as gender may be correspondingly weakened (Brownlow, 1992; Palmer & Peterson, 2016). For example, regardless of changes in the audience' s gender differences, interest in a case, or an expert' s communication skills, professional image is an important predictor of the credibility of expert testimony (Jones et al., 2023). In political communication, audiences' adoption of a speaker' s views depends mainly on the sense of professionalism conveyed by the speaker' s appearance, rather than on the audience' s own political knowledge or cognitive complexity (Palmer & Peterson, 2016). In marketing, a spokesperson' s professional image can still increase brand trust and purchase intention after controlling for product experience or familiarity (Sari et al., 2021).

Accordingly, interviewer image professionalism will moderate the relationship between gender congruence and trust in the interviewer. Interviewer image professionalism reflects applicants' holistic professional judgment based on the external image cues of the AI interviewer; these cues include not only attire, but

also facial features, hairstyle, expressions, and overall visual style. When interviewer image professionalism is high, applicants are likely to form relatively high initial trust, and this "trust anchor" will weaken the effects of other social cues such as gender congruence. Conversely, when interviewer image professionalism is low, applicants lack a clear basis of professionalism and therefore are more likely to rely on gender-congruence cues to judge whether the AI interviewer is reliable, in order to reduce uncertainty and perceived risk. At this point, gender congruence may serve as a compensatory cue and exert a stronger influence on trust. Based on the above analysis, this study proposes the following hypothesis:

Hypothesis 2: Interviewer image professionalism moderates the relationship between AI interviewer–applicant gender congruence and trust in the interviewer; when image professionalism is low, the positive effect of gender congruence on trust in the interviewer is stronger.

1.3 The Mediating Role of Trust in the Interviewer

Applicants' positive experiences during the recruitment process are crucial to organizations (Miles & McCamey, 2018). AI interviews may elicit negative reactions such as feelings of process creepiness and privacy concerns, thereby reducing organizational attractiveness and even causing applicants to interrupt or withdraw from the selection process (Mirowska & Mesnet, 2022). Process creepiness refers to a subjective, contextualized, uncomfortable experience generated by individuals during the AI interview process, typically manifested as unease, a sense of violation, and a sense of threat (Langer et al., 2017). It differs from the classic uncanny valley effect. The latter emphasizes the immediate perceptual conflict triggered by inconsistencies between a humanoid entity' s appearance and behavior (Mori et al., 2012); the former is an overall experience formed toward the entire AI interview process, and may also arise from interaction feedback, decision-making opacity, inexplicability, and inconsistency with social norms (Baek & Kim, 2023; Vimalkumar et al., 2021). Privacy concerns, meanwhile, reflect applicants' perceived risk that personal sensitive information may be collected, misused, or improperly stored (da Motta Veiga & Figueroa-Armijos, 2022; Mori et al., 2025).

Applicants' trust in AI interviewers reflects the extent to which they believe AI can conduct fair, effective, and reliable evaluations (Langer et al., 2023). Higher trust can reduce uncertainty in interactions, leading applicants to interpret the AI interviewer' s behavior as consistent and predictable, thereby reducing discomfort caused by system opacity or inexplicability (Hunkenschroer & Kriebitz, 2023). Therefore, trust is both an evaluation of AI' s capabilities and intentions and a mechanism for buffering uncertainty, making applicants more willing to regard AI interviewers as formal evaluators with professional competence, thus alleviating feelings of process creepiness (Feldkamp et al., 2024). Higher trust also means that applicants believe the AI interview system will comply with privacy-protection norms, restrict personal information to recruitment-evaluation purposes, and adopt necessary storage and access-control measures (Köchling et

al., 2023). Trust also promotes applicants' more positive inferences about the organization's sense of responsibility and goodwill (da Motta Veiga et al., 2023), thereby reducing their concerns about data risks. Accordingly, gender congruence may further reduce process creepiness and privacy concerns by enhancing trust in the interviewer. This study proposes the following hypotheses:

Hypothesis 3a: Interviewer trust mediates the effect of gender congruence on process creepiness. Specifically, gender congruence increases applicants' trust in the interviewer, thereby reducing their process creepiness.

Hypothesis 3b: Interviewer trust mediates the effect of gender congruence on privacy concerns. Specifically, gender congruence increases applicants' trust in the interviewer, thereby reducing their privacy concerns.

In sum, this study proposes a moderated mediation model (see Figure 1): gender congruence between the AI interviewer and the applicant further reduces applicants' process creepiness and privacy concerns by enhancing interviewer trust; meanwhile, interviewer image professionalism moderates the effect of gender congruence on trust. Accordingly, the following hypotheses are proposed:

Hypothesis 4a: Interviewer image professionalism moderates the indirect effect of gender congruence on process creepiness through interviewer trust. When interviewer image professionalism is lower, the negative indirect effect of gender congruence on process creepiness through interviewer trust is stronger.

Hypothesis 4b: Interviewer image professionalism moderates the indirect effect of gender congruence on privacy concerns through interviewer trust. When interviewer image professionalism is lower, the negative indirect effect of gender congruence on privacy concerns through interviewer trust is stronger.

```
flowchart LR
    A["AI interviewer-applicant<br/>gender congruent vs. gender incongruent"] --> B["Intervi
    C["Interviewer image professionalism"] --> B
    B --> D["Privacy concerns"]
    B --> E["Process creepiness"]
```

Figure 1. Research model

1.4 Research Overview

This study tests the above model through two scenario experiments and one experience-sampling study. Study 1 uses a scenario experiment to examine the mediating role of interviewer trust and the moderating effect of interviewer image professionalism; Study 2 uses an experience-sampling method to track applicants' real AI inter—

experience, thereby improving ecological validity. Study 3 used AI interviewers with non-Chinese ethnic appearances, thus testing the robustness of the above mechanism in a cross-cultural interview context.

2 Study 1

2.1 Participants and Design

This experiment adopted a 2 (interviewer gender: male vs. female) × 2 (interviewer image professionalism: high vs. low) between-subjects experimental design. Sample size was estimated using G*Power 3.1 (Faul et al., 2009). With a medium effect size of $f = 0.25$ and a significance level of $\alpha = 0.05$, a two-factor analysis of variance required 171 participants to achieve 90% statistical power. Participants were recruited through the Credamo platform. To ensure that participants had actual interview experience, the recruitment scope was limited to employed individuals. The study used rolling recruitment and excluded invalid questionnaires according to prespecified attention-check criteria, ultimately obtaining 440 valid samples, including 218 men (49.5%) and 222 women (50.5%), with a mean age of 28.92 years ($SD = 3.94$).

2.2 Pre-experiment

Before formal data collection, we conducted a pre-experiment to screen the AI interviewer images to be used in the formal experiment and to verify their validity. Eighty-six participants were recruited on Credamo, including 35 men (40.7%) and 51 women (59.3%), with a mean age of 28.85 years ($SD = 6.36$). We randomly selected 8 male and 8 female image photographs from the Tencent Video platform as the initial materials and presented them to participants in random order. Participants ranked the image professionalism of the male and female interviewers separately. In the data analysis, values were assigned according to the ranking results: the 1st-ranked image was assigned 8 points, decreasing sequentially to 1 point for the 8th-ranked image. Paired-samples t tests showed that, among male images, the male with the highest image professionalism ($M = 7.21$, $SD = 1.10$) differed significantly in professionalism ratings from the male with the lowest image professionalism ($M = 1.98$, $SD = 1.10$), $t(85) = 18.52$, $p < 0.001$, Cohen' s $d = 2.00$. For female images, the ratings of the image with the highest professionalism ($M = 6.91$, $SD = 1.39$) and the image with the lowest professionalism ($M = 2.59$, $SD = 1.47$) likewise differed significantly, $t(85) = 29.16$, $p < 0.001$, Cohen' s $d = 3.15$. These results indicate that AI interviewers with different levels of image professionalism could be clearly distinguished by participants, and thus were suitable as stimulus materials for AI interviewer images in the subsequent experiment. Figure 2a shows the male and female AI interviewer images with high image professionalism, and Figure 2b shows the male and female AI interviewer images with low image professionalism. After determining the AI interviewer image materials, we generated simulated AI interview videos through Tencent Smart Video (https://zenvideo.qq.com/). The simulated videos were all 1 minute and 39 seconds long and covered a complete interview process, including the AI interviewer' s greeting, self-introduction ("Please give a two-minute self-introduction"), interview questions (such as detailed descriptions of strengths and weaknesses, how to improve weaknesses, what

specific measures and experiences were adopted, the greatest challenge encountered in life and how it was handled), and closing segment. To simulate a real interview situation, short pause frames were inserted after each question posed by the AI interviewer to ensure that participants had sufficient time to think and organize their answers. We required participants to answer as they would in a past interview experience, but for privacy protection, they did not need to upload the content of their answers. The AI interview videos can be found at https://osf.io/ezsjp/overview?view_only=7b46317c503d496aa9b514cccb91b8ae.

a High image professionalism

b Low image professionalism

Figure 2. Illustration of AI Interviewers

2.3 Experimental Procedure

In the formal experiment, participants first read a scenario description and were asked to imagine that they were job applicants applying for an ideal position at a company named "Siyuan Company." Next, we informed participants that they would take part in a simulated AI interview for this company in a quiet environment. Participants were then randomly assigned to one of four experimental conditions. To verify whether they had watched the video attentively, participants were required to complete two attention-check questions: first, they were asked to judge whether the AI interviewer they saw was male or female; second, they were asked to identify, from among the options, which interview question had not appeared in the video. Participants who failed to answer correctly were excluded. For the manipulation check, we used one item to measure the professionalism of the interviewer' s image: "The image of the artificial intelligence interviewer looked like that of a professional interviewer," with responses ranging from 1 ("strongly disagree") to 5 ("strongly agree").

Thereafter, participants continued to complete measures of trust in the interviewer, process creepiness, and privacy concerns (1 = "strongly disagree," 5 = "strongly agree"). All English scales were translated into Chinese using a back-translation procedure. The items measuring trust in the interviewer were drawn from Komiak and Benbasat (2006), including six items such as "The artificial intelligence interviewer has strong professional knowledge in evaluation" ; the internal consistency reliability Cronbach' s α was 0.75. Process creepiness was measured using the scale by Langer et al. (2017), including five items such as "During this artificial intelligence interview process, I felt uneasy" ; the internal consistency reliability Cronbach' s α was 0.77. Privacy concerns were measured using the scale by Langer and König (2018), including five items such as "I worry about my privacy" ; the internal consistency reliability Cronbach' s α was

0.87. To control for individual differences in preexisting knowledge of AI, the study measured participants' familiarity with AI (from "1 = not at all familiar" to "5 = very familiar"); the items were adapted from Leo and Huh' s (2020)

study. Finally, participants' demographic information was collected, including gender, age, education level, and whether they worked in an AI-related field.

2.4 Results

Manipulation check. The independent-samples t test showed that the professionalism score in the high image-professionalism group ($M = 4.29$, $SD = 0.60$) was significantly higher than that in the low image-professionalism group ($M = 4.01$, $SD = 0.92$), $t(438) = 3.68$, $p < 0.001$, Cohen' s $d = 0.36$. This indicates that the manipulation of the interviewer' s image professionalism was effective.

A three-factor analysis of variance was conducted with AI interviewer gender, applicant gender, and interviewer image professionalism as independent variables, and interviewer trust as the dependent variable (see Figure 3). The results showed that the three-way interaction among AI interviewer gender, applicant gender, and interviewer image professionalism was significant, $F(1, 432) = 7.86$, $p = 0.005$, $\eta_P^2 = 0.018$. Simple-effects analyses indicated that, for AI interviewers with high image professionalism, when the AI interviewer was male, male applicants' level of trust ($M = 4.41$, $SD = 0.25$) was higher than that of female applicants ($M = 4.17$, $SD = 0.58$), $F(1, 432) = 8.76$, $p = 0.003$, $\eta_P^2 = 0.020$; when the AI interviewer was female, there was no significant difference in trust levels between female and male applicants, $F(1, 432) = 1.34$, $p = 0.248$, $\eta_P^2 = 0.003$. For AI interviewers with low image professionalism, when the AI interviewer was male, male applicants' trust ($M = 4.40$, $SD = 0.39$) was still higher than that of female applicants ($M = 4.07$, $SD = 0.46$), $F(1, 432) = 14.76$, $p < 0.001$, $\eta_P^2 = 0.033$; when the AI interviewer was female, female applicants' trust level ($M = 4.37$, $SD = 0.31$) was higher than that of male applicants ($M = 3.89$, $SD = 0.66$), $F(1, 432) = 31.98$, $p < 0.001$, $\eta_P^2 = 0.069$.

In addition to the three-way interaction, the two-way interaction between AI interviewer gender and applicant gender was significant, $F(1, 432) = 47.04$, $p < 0.001$, $\eta_P^2 = 0.098$. Specifically, male applicants' trust in male AI interviewers ($M = 4.41$, $SD = 0.32$) was significantly higher than that of female applicants ($M = 4.12$, $SD = 0.52$), $F(1, 432) = 23.23$, $p < 0.001$, $\eta_P^2 = 0.051$; whereas female applicants' trust in female AI interviewers ($M = 4.36$, $SD = 0.30$) was significantly higher than that of male applicants ($M = 4.08$, $SD = 0.57$), $F(1, 432) = 23.83$, $p < 0.001$, $\eta_P^2 = 0.052$. The two-way interactions between AI interviewer gender and interviewer image professionalism ($F(1, 432) = 1.92$, $p = 0.167$, $\eta_P^2 = 0.004$), and between applicant gender and interviewer image professionalism were both nonsignificant, $F(1, 432) = 3.25$, $p = 0.072$, $\eta_P^2 = 0.007$. The main effect of AI interviewer gender was also nonsignificant: applicants' trust levels did not differ significantly between male ($M = 4.26$, $SD = 0.46$) and female AI interviewers ($M = 4.22$, $SD = 0.47$), $F(1, 432) = 1.51$, $p = 0.219$, $\eta_P^2 =$

0.003; the main effect of applicant gender was also nonsignificant: male applicants ($M = 4.24$, $SD = 0.49$) and female applicants ($M = 4.24$, $SD = 0.44$)

Figure 3. Effects of AI interviewer gender, applicant gender, and interviewer image professionalism on interviewer trust (Study 1)

Figure 1: Figure 3. Effects of AI interviewer gender, applicant gender, and interviewer image professionalism on interviewer trust (Study 1)

Figure 4. Effects of AI interviewer–applicant gender congruence and interviewer image professionalism on interviewer trust (Study 1)

Figure 2: Figure 4. Effects of AI interviewer–applicant gender congruence and interviewer image professionalism on interviewer trust (Study 1)

showed no significant difference in their level of trust in the AI interviewer, $F(1, 432) = 0.00$, $p = 0.998$, $\eta_P^2 = 0.000$. The main effect of interviewer image professionalism was significant: compared with interviewers low in image professionalism ($M = 4.18$, $SD = 0.51$), applicants trusted interviewers high in image professionalism more ($M = 4.30$, $SD = 0.41$), $F(1, 432) = 7.38$, $p = 0.007$, $\eta_P^2 = 0.017$.

Figure 3. Effects of AI interviewer gender, applicant gender, and interviewer image professionalism on interviewer trust (Study 1)

To further test the research hypotheses, gender congruence (male AI interviewer–male applicant, female AI interviewer–female applicant) was coded as "1," gender incongruence as "0," high interviewer image professionalism as "1," and low interviewer image professionalism as "0." The analysis used gender congruence and interviewer image professionalism as independent variables. The ANOVA results showed that the main effect of AI interviewer–applicant gender congruence was significant: interviewer trust in the gender-congruent group ($M = 4.38$, $SD = 0.31$) was significantly higher than in the gender-incongruent group ($M = 4.10$, $SD = 0.55$), $F(1, 436) = 44.75$, $p < 0.001$, $\eta_P^2 = 0.093$, supporting H1. The main effect of interviewer image professionalism was significant: trust in interviewers with high professionalism ($M = 4.30$, $SD = 0.41$) was significantly higher than in the low-professionalism group ($M = 4.18$, $SD = 0.51$), $F(1, 436) = 7.26$, $p = 0.007$, $\eta_P^2 = 0.016$. The interaction between AI interviewer–applicant gender congruence and interviewer image professionalism was significant, $F(1, 436) = 7.63$, $p = 0.006$, $\eta_P^2 = 0.017$. Simple-effects analysis revealed (see Figure 4) that when interviewer image professionalism was low, gender congruence had a stronger positive effect on trust (gender-congruent group: $M = 4.38$, $SD = 0.35$; gender-incongruent group: $M = 3.98$, $SD = 0.57$; $F(1, 436) = 43.87$, $p < 0.001$, $\eta_P^2 = 0.091$). When interviewer image professionalism was high, gender congruence had a positive effect on ···

the effect on trust was smaller (gender-congruent group: $M = 4.38$, $SD = 0.27$; gender-incongruent group: $M = 4.21$, $SD = 0.50$, $F(1, 436) = 8.35$, $p = 0.004$, $\eta_P^2 = 0.019$), supporting H2.

Figure 4. Effects of AI interviewer–applicant gender congruence and interviewer

image professionalism on interviewer trust (Study 1)

Using R 4.4.0, the Bootstrap method was employed with 5,000 repeated resamples to test the mediating effects. The results showed that the indirect effect of gender congruence on process uncanniness was significant (*indirect effect* = -0.19, *SE* = 0.04, 95% CI [-0.26, -0.12]), and the indirect effect of gender congruence on privacy concerns was significant (*indirect effect* = -0.20, *SE* = 0.04, 95% CI [-0.28, -0.12]), verifying H3a and H3b. A further moderated mediation effect test was conducted using the Bootstrap method, with AI interviewer-applicant gender congruence as the independent variable, interviewer trust as the mediating variable, process uncanniness and privacy concerns as the dependent variables, and interviewer image professionalism as the moderating variable. Age, education level, familiarity with AI, and whether the participant worked in an AI-related field were included as control variables; the results are shown in Table 1. The results indicated that interviewer image professionalism played a moderating role in both paths by which AI interviewer–applicant gender congruence influenced process uncanniness and privacy concerns through interviewer trust (the differences in indirect effects were 0.29 and 0.30, respectively; neither 95% confidence interval included 0). The results suggest that interviewer trust is the mediating mechanism through which gender congruence reduces applicants' process uncanniness and privacy concerns, and that interviewer image professionalism moderates this mediating process, supporting H4a and H4b.

Table 1. Results of the moderated mediation effect test (Study 1)

Path	Interviewer Image Professionalism	Indirect Effect	95% Confidence Interval
Gender congruence → Interviewer trust → Process uncanniness	Low	-0.41	[-0.65, -0.20]
Gender congruence → Interviewer trust → Process uncanniness	High	-0.12	[-0.21, -0.04]
Gender congruence → Interviewer trust → Process uncanniness	Difference (Δ)	0.29	[0.05, 0.56]

Path	Interviewer Image Professionalism	Indirect Effect	95% Confidence Interval
Gender congruence → Interviewer trust → Privacy concerns	Low	-0.42	[-0.69, -0.20]
	High	-0.12	[-0.21, -0.04]
	Difference (Δ)	0.30	[0.04, 0.59]

2.5 Discussion

Study 1 preliminarily verified that gender congruence between the AI interviewer and the applicant can increase applicants' trust in the AI interviewer and further reduce applicants' sense of process strangeness and privacy concerns. At the same time, the results of the analysis of variance showed that the main effects of AI interviewer gender and applicant gender were both nonsignificant; that is, neither gender in itself directly influenced trust in the interviewer. These results indicate that gender itself does not directly affect applicants' trust evaluations of AI interviewers; its effect is mainly reflected in the influence of the congruent matching between AI interviewer gender and applicant gender on trust judgments.

Meanwhile, by manipulating the professionalism of the interviewer image, Study 1 also verified the moderating role of the professionalism of the interviewer image: when the interviewer image was low in professionalism, the indirect effect of gender congruence on process strangeness and privacy concerns through trust in the interviewer was stronger. However, Study 1 used a simulation experiment, which differs from a real job-application scenario. In real recruitment contexts, applicants usually bear a higher level of emotional load because of the real consequences of interview outcomes, and such emotional experiences are often difficult to fully elicit in simulated experimental situations, which may affect their subjective evaluations of and reaction intensity toward the interview process. Therefore, Study 2 further adopted an experience-sampling method to continuously track applicants' experiences interacting with AI interviewers during real job applications, in order to test the robustness and external validity of the effects revealed in Study 1 in a natural setting.

3 Study 2

3.1 Participants and Procedure

Existing AI interview providers in China include Kinlips AI, Niuke AI, AI Yimian, Duomian AI, Beisen AI, Yonyou Dayi, and others. The AI interviewers in these systems include both male and female images, and there are certain differences in the professionalism conveyed by their hairstyles, facial expressions, and attire (see Appendix). Therefore, this study used an experience-sampling method to track applicants' real AI interview experiences, so as to test the applicability of the research model in a natural context. This study recruited participants through snowball sampling via the researchers' social networks. The participants were university students from five colleges and universities who were taking part in autumn recruitment, with a total of 120 participants. Subsequently, a research assistant contacted the participants one by one via WeChat. To protect participants' privacy, their WeChat names were set as participant codes (a combination of the initials of their names and the last digits of their mobile phone numbers). This code also served as the exclusive password for questionnaire completion and was used for questionnaire completion and subsequent data matching. All questionnaires were distributed through the Wenjuanxing platform, and the research assistant sent questionnaire links via WeChat; participants completed and submitted the questionnaires after entering their exclusive questionnaire passwords. The questionnaire survey included an initial baseline questionnaire and daily questionnaires administered for 30 consecutive days. One week before the daily survey began, the baseline questionnaire was distributed to the 120 participants, inviting them to report demographic characteristics such as gender, age, educational level, familiarity with AI, and whether they worked in an AI-related field; 112 valid questionnaires were recovered (response rate = 93.3%). During the subsequent 30 consecutive days, the researchers sampled and investigated participants' AI interview experiences on the day they occurred. The researchers distributed the questionnaire to participants every day at 19:00. To ensure authenticity, at the beginning of the questionnaire

at the time, participants uploaded the interview invitation notification to verify the authenticity of their participation in the AI interview on that day. To ensure that participants' personal privacy would not be disclosed, before uploading the screenshot of the AI interview notification, participants could use a mosaic effect to conceal or blur their names and important personal information. If the screenshot information was incomplete or its authenticity could not be confirmed, the research assistant would further verify it via WeChat. Subsequently, participants were required to answer open-ended questions about the AI interview they had participated in that day, including the interviewing company, interview duration, characteristics of the AI interviewer, and the AI interview technology platform used. Finally, the questionnaire measured the AI interviewer' s gender, the professionalism of the interviewer' s image, trust

in the interviewer, as well as process creepiness and privacy concerns. If participants did not take part in an AI interview that day, they did not need to complete the questionnaire that day. Participants received a reward of 5 yuan each time they completed the questionnaire. After excluding participants who participated only once and invalid data, the final valid sample comprised 1,246 questionnaires from 87 participants, including 39 men (44.8%) and 48 women (55.2%), with a mean age of 23.03 years (SD = 4.08). The number of questionnaires completed by each participant ranged from 2 to 30, with a mean of 14.32, a median of 14, and a standard deviation of 7.72, reflecting questionnaire participation over the entire 30-day follow-up period.

3.2 Measures

All scales in Study 2 were 5-point Likert scales. Trust in the interviewer, privacy concerns, and process creepiness all used the scales employed in Study 1; in Study 2, their α values were 0.65, 0.67, and 0.88, respectively. The professionalism of the interviewer' s image was measured with two items: "How much do you think the AI' s image resembles a professional interviewer?" and "How much do you think the AI interviewer' s attire characteristics resemble those of a professional interviewer?" At the between-person level, this study controlled for age, educational attainment, familiarity with AI, and whether the participant worked in an AI-related field. At the within-person level, considering that experience-sampling data may show trends over time, researchers usually need to control for temporal factors in the research process to reduce the interference of time structure in estimating variable relationships (West & Hepworth, 1991; Trougakos et al., 2014). Therefore, this study controlled for "study day" (the number of days after the start of the study) to exclude possible linear time trends during the one-month follow-up period. In addition, this study adopted an event-triggered follow-up design from experience-sampling research: participants did not complete the questionnaire at a fixed time each day, but only after actually participating in an AI interview on that day (Beal & Weiss, 2003). Because questionnaire completion in this study was triggered by AI interview events, each measurement occurred on a particular day during the one-month follow-up period and also corresponded to the participant' s nth participation in an AI interview during the study period. Therefore, this study further controlled for "interview frequency" (the participant' s nth completion of the questionnaire after completing an AI interview during the study period), in order to minimize interference in psychological responses that might result from familiarization, habituation, or fatigue effects caused by repeated participation in AI interviews and repeated questionnaire completion.

3.3 Analytic Strategy

Considering that the data exhibited a multilevel nested structure (daily data nested within individuals), this study used Mplus 8.7 software (Muthén & Muthén, 2010) to conduct multilevel path analysis. Following the method of

Liu Dong et al. (2018), AI interviewer–applicant gender congruence, interviewer trust, privacy concerns, process creepiness, interviewer image professionalism, and within-person control variables were specified at the within-person level, and between-person control variables were specified at the between-person level. Before hypothesis testing, the within-person predictor variables were group-mean centered

···variables were grand-mean centered. To test the hypothesized moderated mediation effect, Monte Carlo simulations with 20,000 repeated resamples were conducted using R 4.4.0 (Preacher & Selig, 2012), and the statistical significance of the indirect effects was judged using 95% CIs.

3.4 Results

The premise for conducting within-person multilevel analysis is that variables measured daily have sufficient within-person variance. Therefore, this study tested the within-person and between-person variance of the relevant variables, as well as the percentage of within-person variance, by estimating null models (Podsakoff et al., 2019). As shown in Table 2, AI interviewer–applicant gender congruence, interviewer image professionalism, interviewer trust, and process creepiness all showed significant within-person variance, satisfying the prerequisites for multilevel path analysis. By contrast, the within-person variance of privacy concerns did not reach statistical significance; thus, the existing data are insufficient to support stable variation in privacy concerns across different AI interview events. Based on this result, privacy concerns were not included in the subsequent within-person path analysis.

Table 2 Percentage of within-person variance

Variable	Within-person variance (σ^2)	Between-person variance (τ_{00})	Percentage of within-person variance (%)
Gender congruence	0.12**	0.11**	52.17%
Image professionalism	0.13**	0.18**	41.94%
Interviewer trust	0.03**	0.32**	8.57%
Process creepiness	0.09**	0.51**	15.00%
Privacy concerns	0.51	0.35**	59.30%

Note: Percentage of within-person variance = within-person variance / (within-person variance + between-person variance). Gender congruence: AI inter-

viewer-applicant gender congruence = 1; AI interviewer-applicant gender incongruence = 0. *$p < 0.05$, **$p < 0.01$.

Multilevel confirmatory factor analysis was used to test the internal structural validity of the core variables in the model. Because gender congruence is a categorical variable, it was not considered in the multilevel confirmatory factor analysis. The results of the multilevel confirmatory factor analysis showed that the hypothesized three-factor model (interviewer trust, image professionalism, and process creepiness) had a good model fit: $\chi^2 = 641.95$, $df = 121$, $\chi^2/df = 5.31$, CFI = 0.76, TLI = 0.70, RMSEA = 0.06, $\text{SRMR}_{\text{within}} = 0.06$, $\text{SRMR}_{\text{between}} = 0.08$.

To enhance the rigor of the research conclusions, this study used the "controlling for the effects of an unmeasured latent method factor" approach to test for common method bias (Zhou Hao, Long Lirong, 2004). The fit indices of the single-method latent factor model ($\chi^2 = 405.12$, $df = 97$, $\chi^2/df = 4.18$, CFI = 0.86, TLI = 0.77, RMSEA = 0.05, $\text{SRMR}_{\text{within}} = 0.05$, $\text{SRMR}_{\text{between}} = 0.06$), compared with the hypothesized three-factor model, showed that the increases in CFI and TLI did not exceed 0.1, and the decreases in RMSEA, $\text{SRMR}_{\text{within}}$, and $\text{SRMR}_{\text{between}}$ did not exceed 0.05, falling within the reasonable range recommended by Wen Zhonglin et al. (2018). In summary, there was no serious common method bias in this study.

Table 3 presents the means, standard deviations, and correlation coefficients of the variables. At the within-person level, gender congruence was significantly positively correlated with interviewer trust ($r = 0.13$, $p < 0.001$), and interviewer trust was significantly negatively correlated with process creepiness ($r = -0.13$, $p < 0.001$).

Table 3 Matrix of Means, Standard Deviations, and Correlation Coefficients

Variable	1.	2.	3.	4.	5.	6.	7.	8.	9.	10.	11.	12.
Within-person level												
1. AI interviewer gender	-	-0.07	0.13	0.21	-0.19	0.16	0.21	0.43**	-0.06	-0.12	-0.03	-0.10

Variable	1.	2.	3.	4.	5.	6.	7.	8.	9.	10.	11.	12.
2. Gender congruence	0.14**	- 0.05	-	0.02	0.02	- 0.15	- 0.09	- 0.70**	- 0.07	- 0.06	0.19	- 0.06
3. Image professionalism	- 0.04	- 0.03	-	0.77**	- 0.35**	- 0.13	0.42**	0.13	0.07	- 0.43**	0.17	- 0.20
4. Interviewer trust	0.04	0.13**	0.25**	(0.65)	- 0.52**	0.03	0.50**	0.12	0.03	- 0.50**	0.20	- 0.23*
5. Process creepiness	0.01	- 0.01	0.04	- 0.13**	(0.88)	0.01	- 0.28**	- 0.03	- 0.03	0.28**	- 0.03	0.21*
6. Study days	- 0.02	- 0.01	0.09**	0.03	- 0.04	-	0.38**	0.09	0.07	0.08	- 0.14	0.33**
7. Number of interviews	- 0.02	- 0.01	0.08**	0.02	- 0.03	0.97**	-	0.13	0.06	- 0.46**	- 0.03	0.04
Between-person level												
8. Applicant gender								-	- 0.05	- 0.03	- 0.16	- 0.13

Variable	1.	2.	3.	4.	5.	6.	7.	8.	9.	10.	11.	12.
9. Age									-	- 0.09	0.03	- 0.09
10. Education level										-	- 0.01	0.08
11. Familiarity with AI											-	- 0.40**
12. Whether worked in an AI-related field												-
Mean	0.27	0.63	4.37	4.38	1.64	16.91	9.72	0.48	23.23	2.98	4.04	1.33
Standard deviation$_{\text{within-person}}$	0.45	0.48	0.48	0.37	0.69	7.53	6.45					
Standard deviation$_{\text{between-person}}$	0.29	0.35	0.33	0.33	0.63	3.40	3.16	0.50	2.89	0.36	0.72	0.47

Note: $N_{\text{within-person}} = 1246$; $N_{\text{between-person}} = 87$. Correlations below the diagonal are within-person correlation coefficients; correlations above the diagonal are between-person correlation coefficients. Values in parentheses on the diagonal are the Cronbach' s α reliability coefficients of the variables. AI interviewer gender: male = 1, female = 0; gender congruence: AI interviewer–applicant gender congruent = 1, AI interviewer–applicant gender incongruent = 0; applicant gender: male = 1, female = 0; education level: high school or below = 1, junior college = 2, bachelor' s degree = 3, master' s degree or above = 4; whether

worked in an AI-related field: yes = 1, no = 2. $^{*}p < 0.05$, $^{**}p < 0.01$.

Table 4 reports the results of the multilevel path analysis. Gender congruence significantly and positively predicted interviewer trust ($\gamma = 0.10$, $SE = 0.03$, $p < 0.001$), supporting H1. Interviewer trust significantly and negatively predicted process creepiness ($\gamma = -0.26$, $SE = 0.09$, $p < 0.001$). Using R 4.4.0 to conduct a Monte Carlo simulation with 20,000 repeated draws, the results showed that gender congruence reduced process creepiness through increasing interviewer trust, and this indirect effect was significant (*indirect effect* = -0.015, 95% CI [-0.035, -0.001]); thus, H3a was supported.

Table 4 Results of Multilevel Path Analysis

Variable	Interviewer Trust	Interviewer Trust	Process Creepiness	Process Creepiness
	Estimate	*S.E.*	*Estimate*	*S.E.*
Within-person level				
Study day	-0.003	0.01	-0.01	0.01
Interview session	0.01	0.01	0.01	0.01
Gender congruence	0.10**	0.03	0.01	0.02
Image professionalism	0.12**	0.03	0.06	0.03
Interviewer trust			-0.26**	0.09
Interaction term				
Gender congruence × image professionalism	-0.27**	0.10	-0.06	0.07
Between-person level				
Age	-0.02	0.02	0.003	0.02
Educational level	-0.27*	0.12	0.30	0.14
Familiarity with AI	-0.01	0.05	0.10	0.14

Figure 5. The moderating effect of interviewer image professionalism on the relationship between gender congruence and trust in the interviewer

Figure 3: Figure 5. The moderating effect of interviewer image professionalism on the relationship between gender congruence and trust in the interviewer

Variable	Interviewer Trust	Interviewer Trust	Process Creepiness	Process Creepiness
Whether working in an AI-related field	-0.05	0.08	0.32	0.19

Note: Gender congruence: AI interviewer–applicant gender congruent = 1, AI interviewer–applicant gender incongruent = 0; educational level: high school or below = 1, junior college = 2, bachelor' s degree = 3, master' s degree or above = 4; whether working in an AI-related field: yes = 1, no = 2. * $p < 0.05$, ** $p < 0.01$.

As shown in Table 4, the interaction term between gender congruence and image professionalism had a significant negative effect on interviewer trust ($\gamma = -0.27$, $SE = 0.10$, $p < 0.001$), and the moderation effect plot is shown in Figure 5. Simple slope tests indicated that when interviewer image professionalism was low, the relationship between gender congruence and interviewer trust was significantly positive (*slope* = 0.19, $p < 0.001$); when interviewer image professionalism was high, there was no significant relationship between gender congruence and interviewer trust (*slope* = 0.01, $p = 0.711$). Overall, H2 was supported. Using R 4.4.0 to conduct a Monte Carlo simulation with 20,000 repeated draws, the results showed that when interviewer image professionalism was low, gender congruence reduced process creepiness by increasing interviewer trust, and this indirect effect was significant (*indirect effect* = -0.035, 95% CI [-0.064, -0.012]); however, when interviewer image professionalism was high, the above indirect effect was not significant (*indirect effect* = -0.004, 95% CI [-0.020, 0.010]). In addition, the difference between the indirect effects under the high and low conditions was significant (*indirect effect* = 0.031, 95% CI [0.007, 0.063]). Therefore, H4a was supported.

Legend: dashed line with diamond markers = low image professionalism; solid line with square markers = high image professionalism. X-axis: gender incongruence; gender congruence.

Figure 5. The moderating effect of interviewer image professionalism on the relationship between gender congruence and trust in the interviewer

3.5 Discussion

Study 2 further verified that gender congruence between the AI interviewer and the applicant can enhance trust in the interviewer and, through interviewer trust, reduce applicants' sense of process creepiness; at the same time, it again supported the moderating role of interviewer image professionalism. In addition, as shown in Table 3, at the within-person level, there was no significant correlation between AI interviewer gender and interviewer trust ($r = 0.04, p = 0.179$); at the between-person level, the applicant' s gender likewise showed no significant correlation with interviewer trust ($r = 0.12, p = 0.367$). These results indicate that no significant direct relationship was found between either AI interviewer gender or applicant gender and interviewer trust; the role of gender cues is mainly manifested as a congruence effect between the two.

However, in the present study, the within-person variance in privacy concerns did not reach a significant level, indicating that applicants' privacy concerns were relatively stable across different AI interview contexts. This result may stem from two reasons. First, this study used an experience-sampling method and measured applicants' retrospective evaluations of AI interviews in which they had actually participated. One possible explanation is that applicants who actually took part in AI interviews had already shown a certain level of acceptance of relevant privacy risks before the interview, thereby limiting event-level fluctuations in privacy concerns. Second, the importance of privacy issues may differ across cultural contexts. Prior research has indicated that, compared with Western countries, Chinese applicants are generally less sensitive to privacy risks in technological applications (Xing et al., 2023). The above results suggest that the variability of privacy concerns across different contexts may be influenced by cultural background. Further, this also raises a more important question: when AI interviewers present external images of different races, whether the mechanism by which gender congruence influences applicant reactions through interviewer trust, as well as the moderating effect of interviewer image professionalism, still exists stably. On this basis, Study 3 further extends the context to AI interviewers with different racial appearances, focusing on whether, under conditions in which cultural and appearance features are incongruent, the mediating role of interviewer trust and the moderating role of interviewer image professionalism remain robust, thereby enhancing the external validity of the research conclusions across different recruitment contexts.

4 Study 3

Against the backdrop of the increasing prevalence of globalized recruitment and foreign-language talent selection, more and more organizations are adopting English AI interview systems (e.g., HireVue, Talview, Vervoe, Modern Hire, etc.). In such contexts, applicants often need to interact with AI interviewers whose racial/ethnic appearance differs from their own. Therefore, Study 3 no longer used AI interviewer appearances based on Chinese cultural backgrounds, as in Studies 1 and 2, but instead adopted AI interviewers from other ethnic groups to

simulate a more representative cross-cultural interview context in international recruitment processes, thereby further enhancing the external validity of the research. Overall, this study continued the experimental paradigm of Study 1, but made several optimizations in the experimental design to improve internal validity. First, with respect to the manipulation of interviewer appearance, in Study 1 the high- and low-professionalism appearance groups differed in attire (formal wear vs. casual wear), and this difference may have been the main source influencing perceptions of professionalism. However, in real interview contexts, interviewers usually appear in formal attire. Therefore, Study 3 uniformly used formally dressed images and, on this basis, manipulated the professionalism of the appearance, so as to reduce potential confounding caused by differences in clothing. Second, although the manipulation check results in Study 1 indicated that the manipulation of appearance professionalism was effective, it may still have been accompanied by changes in other potential confounding variables (such as AI interview technology credibility, organizational cultural tightness, and organizational credibility), thereby affecting causal inference. For this reason, while conducting the manipulation check for appearance professionalism, Study 3 additionally measured the above variables to test whether there were differences between experimental groups, thereby ruling out the influence of potential confounds. Third, although this article conceptually distinguishes process creepiness from the "uncanny valley effect," the former refers to applicants' overall negative emotional experience during the AI interview process, whereas the latter refers to individuals' immediate perceptual conflict and discomfort in response to humanoid AI interviewer images (i.e., perceived AI interviewer uncanniness). However, perceived AI interviewer uncanniness may serve as an important antecedent variable influencing interviewer trust, process creepiness, and privacy concerns. Therefore, in the model testing of Study 3, perceived AI interviewer uncanniness was included in the analysis as a control variable.

4.1 Participants and Design

This experiment adopted a 2 (interviewer gender: male vs. female) × 2 (interviewer appearance professionalism: high vs. low) between-subjects experimental design. Sample size was estimated using G*Power 3.1 software (Faul et al., 2009). Assuming a medium effect size of $f = 0.25$ and a significance level of $\alpha = 0.05$, a two-factor analysis of variance required 171 participants to achieve 90% statistical power. In this study, a total of 331 valid participants were recruited through the Credamo platform. All participants were employed individuals and had real job-interview experience. Among them, 144 were male (43.5%) and 187 were female (56.5%); the mean age was 30.82 years ($SD = 7.89$).

4.2 Pre-experiment

Before formal data collection, we conducted a pre-experiment to screen the AI interviewer images to be used in Study 3 and to test their validity. Based on the U.S. Colossyan platform (https://www.colossyan.com/), we first screened

Examples of U.S. AI interviewers: high image professionalism and low image professionalism

Figure 4: Examples of U.S. AI interviewers: high image professionalism and low image professionalism

AI images wearing formal attire and, on this basis, selected four male and four female images, covering typical U.S. ethnic groups such as White and Latino (Contreras et al., 2024). Through the Credamo platform, we recruited 120 participants, including 36 males (30.0%) and 84 females (70.0%), with a mean age of 30.27 years ($SD = 8.33$). Participants separately ranked the professionalism of the male and female images. In the data analysis, according to

The ranking results were scored: rank 1 was scored as 4 points, and rank 4 was scored as 1 point. The paired-samples t test results showed that the male image with the highest professionalism ($M = 3.12$, $SD = 1.01$) differed significantly in professionalism ratings from the male image with the lowest professionalism ($M = 1.74$, $SD = 0.85$), $t(119) = 17.85$, $p < 0.001$, Cohen' s $d = 1.63$. For female images, ratings likewise differed significantly between the image with the highest professionalism ($M = 3.55$, $SD = 0.74$) and the image with the lowest professionalism ($M = 1.51$, $SD = 0.79$), $t(119) = 10.31$, $p < 0.001$, Cohen' s $d = 0.94$. These results indicate that AI interviewer images with different levels of professionalism can be clearly distinguished by participants and are therefore suitable as stimulus materials for the AI interviewer images in the subsequent experiment. Figure 6a presents the male and female AI interviewer images with high image professionalism, and Figure 6b presents the male and female AI interviewer images with low image professionalism. After determining the AI interviewer images, we used the Colossyan platform to convert these static images into AI interview videos. The video content remained structurally consistent with Study 1, but the AI interviewers used English expressions. The specific videos can be found at osf(https://osf.io/ezsjp/overview?view_only=7b46317c503d496aa9b514cccb91b8ae).

a High image professionalism

b Low image professionalism

Figure 6 Examples of U.S. AI interviewers

4.3 Experimental Procedure

In the formal experiment, all participants were randomly assigned to one of the four experimental conditions. The overall procedure of Study 3 was basically consistent with that of Study 1. To ensure that participants watched the simulated video content carefully, they were required to answer two attention-check questions. Subsequently, participants completed the manipulation-check items for image professionalism, including "The artificial-intelligence interviewer' s

image looks like that of a professional interviewer" and "The overall image of the artificial-intelligence interviewer' s ···"

reflect a relatively high level of professionalism," and "In terms of external image, the artificial intelligence interviewer is professional" ; Cronbach' s α was 0.79. The response scale ranged from 1 ("strongly disagree") to 5 ("strongly agree").

To examine the discriminant validity of the manipulation and to control for potential confounding variables, participants further completed measures of AI interview technology credibility, organizational cultural tightness, and organizational credibility (1 = "strongly disagree," 5 = "strongly agree"). Among them, the AI interview technology credibility scale was adapted from Chowdhury et al. (2022), including six items such as "I have confidence in the application of artificial intelligence interview technology" ; Cronbach' s α was 0.66. The items measuring organizational cultural tightness were adapted from Lim and Chua (2026), including five items such as "There are many rules, regulations, and behavioral norms in the organization (Siyuan Company) that employees need to follow" ; Cronbach' s α was 0.78. The items measuring organizational credibility were adapted from Newell and Goldsmith (2001), including four items such as "I trust this organization (Siyuan Company)" ; Cronbach' s α was 0.67.

Subsequently, participants completed the AI interviewer eeriness perception scale, the interviewer trust scale (same as Study 1, $\alpha = 0.72$), the process creepiness scale (same as Study 1, $\alpha = 0.73$), and the privacy concerns scale (same as Study 1, $\alpha = 0.87$). Perceived eeriness of the AI interviewer was measured using the Ho and MacDorman (2010) scale. Participants evaluated the overall image of the AI interviewer in this interview (including appearance, facial expression, and voice); items included eight items such as "reassuring–eerie" and "natural–eerie," with Cronbach' s $\alpha = 0.77$.

Finally, participants reported their demographic information, including gender, age, educational level, familiarity with AI, and whether they worked in an AI-related field.

4.4 Results

Manipulation check. The results of an independent-samples t test showed that professionalism ratings in the high-professionalism interviewer image condition ($M = 4.46$, $SD = 0.26$) were significantly higher than those in the low-professionalism interviewer image condition ($M = 4.24$, $SD = 0.55$), $t(329) = 4.81$, $p < 0.001$, Cohen' s $d = 0.53$. In addition, the differences between the high- and low-image-professionalism groups on other variables were not significant, including AI interview technology credibility ($t(329) = 1.84$, $p = 0.067$, Cohen' s $d = 0.20$), organizational cultural tightness ($t(329) = 1.74$, $p = 0.082$, Cohen' s $d = 0.19$), and organizational credibility ($t(329) = 1.75$, $p = 0.082$, Cohen' s $d = 0.19$). The above results indicate that the manipulation of interviewer

image professionalism was effective and affected only participants' perceptions of interviewer professionalism.

The results of the three-way analysis of variance showed that the three-way interaction among AI interviewer gender, applicant gender, and interviewer image professionalism was significant, $F(1, 323) = 8.44$, $p = 0.004$, $\eta_P^2 = 0.025$. Simple effects analyses showed (see Figure 7) that, for AI interviewers with high image professionalism, when the AI interviewer was male, there was no significant difference in trust levels between female and male applicants, $F(1, 323) = 3.77$, $p = 0.053$, $\eta_P^2 = 0.012$; when the AI interviewer was female, there was no significant difference in trust levels between female and male applicants, $F(1, 323) = 1.89$, $p = 0.171$, $\eta_P^2 = 0.006$. For AI interviewers with low image professionalism, when the AI interviewer was male, male applicants ($M = 4.35$, $SD = 0.25$) showed significantly higher trust than female applicants ($M = 4.00$, $SD = 0.53$), $F(1, 323)$

$= 19.92$, $p < 0.001$, $\eta_P^2 = 0.058$; when the AI interviewer was female, there was a marginally significant difference in the level of trust between female and male applicants, $F(1, 323) = 3.61$, $p = 0.058$, $\eta_P^2 = 0.011$.

The two-way interaction between AI interviewer gender and applicant gender was significant, $F(1, 323) = 12.31$, $p < 0.001$, $\eta_P^2 = 0.037$. Specifically, male applicants' trust in male AI interviewers ($M = 4.37$, $SD = 0.24$) was significantly higher than that of female applicants ($M = 4.12$, $SD = 0.45$), $F(1, 323) = 20.50$, $p < 0.001$, $\eta_P^2 = 0.060$; whereas under the female AI interviewer condition, applicant gender did not significantly affect trust, and there was no significant difference between female applicants ($M = 4.22$, $SD = 0.30$) and male applicants ($M = 4.20$, $SD = 0.44$), $F(1, 323) = 0.21$, $p = 0.646$, $\eta_P^2 = 0.001$. The two-way interaction between AI interviewer gender and interviewer image professionalism ($F(1, 323) = 0.14$, $p = 0.707$, $\eta_P^2 = 0.000$), as well as the two-way interaction between applicant gender and interviewer image professionalism, was not significant, $F(1, 323) = 0.17$, $p = 0.685$, $\eta_P^2 = 0.001$. The main effect of AI interviewer gender was not significant: applicants' trust levels did not differ significantly between male ($M = 4.22$, $SD = 0.40$) and female AI interviewers ($M = 4.21$, $SD = 0.37$), $F(1, 323) = 0.66$, $p = 0.416$, $\eta_P^2 = 0.002$. In contrast, the main effect of applicant gender was significant: male applicants ($M = 4.28$, $SD = 0.37$) had significantly higher trust in AI interviewers than female applicants ($M = 4.16$, $SD = 0.39$), $F(1, 323) = 8.15$, $p = 0.005$, $\eta_P^2 = 0.025$; the main effect of interviewer image professionalism was also significant: compared with interviewers with low image professionalism ($M = 4.15$, $SD = 0.46$), applicants trusted interviewers with high image professionalism more ($M = 4.28$, $SD = 0.29$), $F(1, 323) = 9.85$, $p = 0.002$, $\eta_P^2 = 0.030$.

Figure 7. Effects of AI interviewer gender, applicant gender, and interviewer image professionalism on trust in the interviewer (Study 3)

To further test the research hypotheses, we used gender congruence and interviewer image professionalism as independent variables. The results of the anal-

Bar chart showing interviewer trust by low vs. high image professionalism and four gender combinations. Y-axis: interviewer trust. Low image professionalism: male AI interviewer–male applicant 4.35; female AI interviewer–female applicant 4.23; male AI interviewer–female applicant 4.00; female AI interviewer–male applicant 4.07. High image professionalism: male AI interviewer–male applicant 4.39; female AI interviewer–female applicant 4.21; male AI interviewer–female applicant 4.24; female AI interviewer–male applicant 4.32.

Figure 5: Bar chart showing interviewer trust by low vs. high image professionalism and four gender combinations. Y-axis: interviewer trust. Low image professionalism: male AI interviewer–male applicant 4.35; female AI interviewer–female applicant 4.23; male AI interviewer–female applicant 4.00; female AI interviewer–male applicant 4.07. High image professionalism: male AI interviewer–male applicant 4.39; female AI interviewer–female applicant 4.21; male AI interviewer–female applicant 4.24; female AI interviewer–male applicant 4.32.

ysis of variance showed that the main effect of AI interviewer–applicant gender congruence was significant: trust in the interviewer was significantly higher in the gender-congruent group ($M = 4.29,\ SD = 0.28$) than in the gender-incongruent group ($M = 4.15,\ SD = 0.45$), $F(1, 327) = 11.82,\ p < 0.001,\ \eta_P^2 = 0.035$; the main effect of interviewer image professionalism was significant: trust in interviewers with high image professionalism ($M = 4.28,\ SD = 0.29$) was significantly higher than in the low-professionalism group ($M = 4.15,\ SD = 0.46$), $F(1, 327) = 8.70,\ p = 0.003,\ \eta_P^2 = 0.026$; and the interaction between AI interviewer–applicant gender congruence and interviewer image professionalism was significant, $F(1, 327) = 9.54,\ p = 0.002,\ \eta_P^2 = 0.028$. Simple-effects analysis found (see Figure 8): when interviewer image professionalism was low, gender congruence had a significant positive effect on trust (gender-congruent group: $M = 4.29,\ SD = 0.27$; gender-incongruent group: $M = 4.03,\ SD = 0.55$; $F(1, 327) = 20.34,\ p < 0.001,\ \eta_P^2 = 0.059$); by contrast, when interviewer image professionalism was high, the effect of gender congruence on trust was not significant (gender-congruent group: $M = 4.29,\ SD = 0.28$; gender-incongruent group: $M = 4.27,\ SD = 0.29,\ F(1, 327) = 0.06,\ p = 0.800,\ \eta_P^2 = 0.000$). H2 was supported.

Figure 8. Effects of AI interviewer–applicant gender congruence and interviewer image professionalism on interviewer trust (Study 3)

Using R 4.4.0, the Bootstrap method was employed to conduct 5,000 repeated resampling tests of the mediating effects and the moderated mediating effects. Age, education level, familiarity with AI, whether the participant had worked in an AI-related field, and perceived uncanniness of the AI interviewer were included as control variables. The results indicated that the perceived uncanniness of the AI interviewer did not significantly predict interviewer trust (*effect* = -0.07, $p = 0.135$), but it had a significant positive effect on process uncanni-

Bar chart showing interviewer trust by gender congruence and professionalism: under gender incongruence, low image professionalism = 4.03 and high image professionalism = 4.27; under gender congruence, low image professionalism = 4.29 and high image professionalism = 4.29. Y-axis: interviewer trust. Legend: low image professionalism; high image professionalism.

Figure 6: Bar chart showing interviewer trust by gender congruence and professionalism: under gender incongruence, low image professionalism = 4.03 and high image professionalism = 4.27; under gender congruence, low image professionalism = 4.29 and high image professionalism = 4.29. Y-axis: interviewer trust. Legend: low image professionalism; high image professionalism.

ness (*effect* = 0.21, $p < .001$). Further mediation-effect analyses showed that the indirect effect of gender congruence on process uncanniness was significant (*indirect effect* = -0.10, *SE* = 0.03, 95% CI [-0.17, -0.04]), and the indirect effect of gender congruence on privacy concerns was significant (*indirect effect* = -0.11, *SE* = 0.03, 95% CI [-0.18, -0.05]), thereby supporting H3a and H3b. The results of the moderated mediation analysis showed that interviewer image professionalism played a moderating role in both pathways through which AI interviewer–applicant gender congruence influenced process uncanniness and privacy concerns via interviewer trust (the differences in indirect effects were 0.36 and 0.46, respectively, and the 95% confidence intervals did not include 0). Therefore, H4a and H4b were supported.

Table 5. Results of the Test of Moderated Mediation Effects (Study 3)

Path	AI Interviewer Persona Professionalism	Indirect Effect	95% Confidence Interval
Gender congruence → interviewer trust → perceived process strangeness	Low	-0.37	[-0.60, -0.15]
Gender congruence → interviewer trust → perceived process strangeness	High	-0.01	[-0.07, 0.06]

Path	AI Interviewer Persona Professionalism	Indirect Effect	95% Confidence Interval
Gender congruence → interviewer trust → perceived process strangeness	Difference (Δ)	0.36	[0.12, 0.63]
Gender congruence → interviewer trust → privacy concern	Low	-0.47	[-0.72, -0.22]
Gender congruence → interviewer trust → privacy concern	High	-0.01	[-0.09, 0.07]
Gender congruence → interviewer trust → privacy concern	Difference (Δ)	0.46	[0.17, 0.76]

4.5 Discussion

Study 3, in a context where AI interviewers were presented as having racially different external personas, and through a more rigorous experimental design, again verified the mediating role of interviewer trust and the moderating role of interviewer persona professionalism, indicating that the above mechanism has relatively high robustness in cross-racial contexts. Specifically, in terms of experimental design, Study 3 further optimized the manipulation check for persona professionalism. To test the discriminant validity of the manipulation and rule out potential confounding variables, the study additionally measured AI interview technology credibility, organizational cultural tightness-looseness, and organizational credibility. The results showed that the manipulation of interviewer persona professionalism was effective and affected only applicants' perceptions of interviewer professionalism, without significantly affecting the above variables. This indicates that the manipulation mainly operated on persona professionalism itself rather than on other related cognitive factors.

At the same time, the ANOVA results showed that the main effect of AI interviewer gender was not significant; that is, applicants' level of trust in male

versus female AI interviewers did not differ significantly. However, unlike Study 1, the main effect of applicant gender was significant: male applicants' trust in AI interviewers was significantly higher than that of female applicants. This result is consistent with existing research. For example, Zhang and Yencha (2022), based on a large-scale survey including 4,245 respondents, found that gender has a significant and robust effect on individuals' trust in and acceptance of AI recruitment algorithms. Specifically, after controlling for variables such as race, age, income, education level, perceived algorithmic bias, and trust in technology companies, men scored significantly higher than women in perceived fairness, effectiveness, and overall acceptability, whereas women were more inclined to believe that AI recruitment systems are unfair, less effective, and difficult to accept; moreover, this difference was especially pronounced in AI video interview contexts.

5 General Discussion

Based on similarity-attraction theory, this article, through two experiments and one experience-sampling study, reached the following conclusions: First, AI interviewer–applicant gender congruence can enhance applicants' trust in AI interviewers; second, interviewer persona professionalism moderates this relationship, such that when persona professionalism is low, applicants rely more on gender-congruence cues, and the positive effect of gender congruence on trust is more pronounced, reflecting a compensatory effect; third, interviewer trust is an important psychological mechanism through which gender congruence affects applicant reactions: gender congruence, by increasing inter-

interviewer trust further reduces process creepiness and privacy concerns; fourth, when AI interviewers present different racialized appearances, the above mediating and moderating mechanisms still exist.

5.1 Theoretical Contributions

First, this study extends research on applicant reactions to AI interviews by shifting the focus from technical performance to the external image characteristics of AI interviewers, and by revealing the mediating role of interviewer trust between image cues and applicant reactions. Applicant reactions refer to applicants' attitudes, feelings, or beliefs about the recruitment process (Ryan & Ployhart, 2000), which further influence intention to apply, willingness to accept a job, and perceived organizational attractiveness (McCarthy et al., 2017). However, existing research has mainly discussed the efficiency, fairness, and technical features of AI recruitment, paying insufficient attention to how the external image of AI interviewers can improve the applicant experience (Lukacik et al., 2022). This article identifies two important cues—gender congruence and image professionalism—and finds that they can influence process creepiness and privacy concerns through interviewer trust. This indicates that the external presentation of AI interviewers is not a neutral interface attribute, but rather

an important source of information through which applicants understand AI interviews and form subsequent reactions.

Second, this study deepens and extends the applicability of similarity-attraction theory in AI recruitment contexts. Similarity-attraction theory suggests that people tend to prefer and trust objects that share similar characteristics with themselves (Byrne, 1997), and that perceived similarity plays an important role in interpersonal attraction, relationship building, and trust formation (Chen & Park, 2021). This study finds that when the AI interviewer' s gender is congruent with that of the applicant, applicants trust the AI interviewer more. This suggests that when interacting with AI interviewers, applicants still form perceptions of similarity based on anthropomorphic social cues (Ahn et al., 2022), thereby extending similarity-attraction theory from interpersonal interaction to human–machine interaction in recruitment contexts. It should be noted that the applicability of this theory in human–machine interaction is not without boundaries; its premise is that AI possesses sufficient social cues to activate users' social cognition. Glikson and Woolley (2020) divided AI presentation forms into robotic AI, virtual AI, and embedded AI. Robotic AI and virtual AI can present anthropomorphic characteristics through external appearance, voice, natural language, and role identity, thereby enhancing social presence (Pitardi et al., 2023); embedded AI is usually hidden behind a system or platform, lacking a clear external image and interaction identity, so users are more likely to make judgments based on accuracy, reliability, transparency, explainability, and procedural fairness. AI interviewers are closer to virtual AI, possessing a visualized image, gender cues, voice, and an interviewer role, and therefore can activate applicants' perceptions of similarity and initial trust judgments. Thus, similarity-attraction theory is more applicable to robotic AI and virtual AI with social cues and interactive roles, and should not be unconditionally generalized to all AI systems.

Third, this study extends the boundaries of research on AI anthropomorphism and reveals compensatory mechanisms among different anthropomorphic cues. Existing research has mainly focused on the main effects of a single type of anthropomorphic feature, such as appearance, language, or behavior, with relatively little discussion of how multiple cues jointly influence trust (Zhou Lulu et al., 2025). This study finds that different external image cues are not simply additive: when an AI interviewer presents relatively high image professionalism, this strong cue can lead applicants to form relatively stable trust judgments, and the effects of other cues such as gender congruence are consequently weakened; when image professionalism is relatively low, applicants are more inclined to use available cues such as gender congruence to reduce uncertainty, thereby making their trust in the AI interviewer···

···the effect of trust is strengthened. This compensatory pattern indicates that users dynamically adjust the basis of their social judgments according to the availability and diagnosticity of different cues, rather than having trust determined directly by a single design element. In this way, the present study fills

a gap in the existing literature, which has paid insufficient attention to multi-cue interaction mechanisms (Li & Suh, 2022), and also demonstrates how users dynamically adjust their social judgment strategies in human–AI interaction environments where multiple anthropomorphic cues coexist, thereby providing a new theoretical perspective for understanding the formation of human–AI trust.

5.2 Practical Implications

Applicants' feelings are crucial for organizations to attract and retain talent. This study reveals the influence of AI interviewers' external images on applicant reactions, and also provides practical implications for organizations in building and deploying AI interview systems.

First, gender congruence between the AI interviewer and the applicant is a readily achievable strategy for enhancing trust. This study found that when the AI interviewer' s gender matches that of the applicant, applicants show higher trust in the interviewer. This suggests that the external presentation of an AI interviewer is not a neutral technical variable, but an important design cue that influences applicants' experiences. As applicants increasingly come to view AI interviewers as symbolic representatives of organizations, merely improving algorithmic quality and operational efficiency may be insufficient to improve the interview experience. By contrast, gender matching, as an easily implemented adjustment of interface features, can enhance applicants' sense of closeness and perceived credibility at very low cost, and reduce discomfort caused by technology-mediated interaction. Therefore, when selecting or developing AI interview systems, organizations may provide diverse and user-selectable AI interviewer images so as to establish a positive psychological foundation at the initial stage of the interview, strengthen candidates' overall favorable impressions of the organization, and enhance attractiveness in a highly competitive talent market.

Second, the image professionalism of AI interviewers provides a new entry point for organizations to optimize interface design and interaction experience. When image professionalism is relatively low, applicants rely more on social cues such as gender congruence to judge whether the AI interviewer is trustworthy; when image professionalism is relatively high, this reliance is correspondingly weakened. This indicates that the maturity of facial features, the stability of expressions and eye gaze, the appropriateness of attire, and the fit between the overall image and the position and industry context all affect applicants' judgments of AI interviewers' professionalism. Higher image professionalism helps AI interviewers to be regarded as reliable, competent evaluators who deserve to be taken seriously. For system developers, while optimizing the underlying algorithms, moderately adjusting the external image of AI interviewers may also improve applicants' experiences at relatively low cost.

Finally, trust in the interviewer is an important psychological mechanism connecting the external image of the interviewer with applicant reactions. Gender

congruence can increase applicants' trust in AI interviewers, thereby further reducing feelings of process creepiness and privacy concerns. This suggests that when promoting AI recruitment, organizations should not evaluate technological applications solely from the perspectives of efficiency, cost, and scalability, but should also attend to applicants' emotional experiences and perceptions of risk. While optimizing the image of AI interviewers, organizations should also improve process transparency; clearly explain the evaluation methods as well as rules for data collection, use, and storage; and establish necessary data governance and privacy protection measures. By creating a trustworthy AI interview environment, organizations can alleviate candidates' unease about technology-mediated interaction in a more humanized way, thereby improving interview acceptance and organizational attractiveness.

5.3 Research Limitations and Future Directions

This study has the following limitations. First, Studies 1 and 3 used online scenario experiments, in which participants watched simulated interview videos; this still differs from the real consequences involved in actual recruitment interactions (Aguinis & Bradley, 2014). Although this video-based presentation is consistent with the typical current process of AI interviews (Lavanchy et al., 2023), real job applicants may experience stronger anxiety and engagement because of the importance of the outcomes, and thus may respond differently to AI interviewer cues. Future research could adopt field experiments, laboratory experiments, or semi-natural settings, allowing participants to interact with AI interviewers in environments closer to real job seeking, thereby improving ecological validity and strengthening control over the interaction process and alternative explanations. In addition, although this study combined experimental methods with experience-sampling methods, the mediator and outcome variables were measured primarily at the same time point. Future research could use time-lagged or longitudinal designs to further examine the causal order among variables.

Second, Study 2 distributed questionnaires at a fixed time each day, which may have missed other AI interviews that participants experienced on the same day. Although this approach is relatively common in experience-sampling research (Foulk & Lanaj, 2022), future studies could still measure immediately after each interview to obtain more fine-grained process data. In addition, repeated measurement within one month may produce familiarization, habituation, or fatigue effects (Beal & Weiss, 2003). Although this study controlled for study days and number of interviews in the analyses to minimize the influence of temporal trends, measurement order, and repeated exposure on the research results, this treatment still cannot completely rule out fatigue and habituation interference caused by long-term repeated participation in AI interview tasks. Future research could compare psychological responses to first-time versus repeated exposure to AI interviewers, or appropriately reduce the density of repeated measurements.

Third, all participants in this study were Chinese job applicants, which may limit the cross-cultural applicability of the research conclusions. Cultural groups may differ in their attitudes toward artificial intelligence, technological risks, and privacy protection. For example, some research has indicated that the Chinese public overall has a relatively higher level of trust in artificial intelligence (Gillespie et al., 2023). This cultural characteristic may influence candidates' trust judgments and risk perceptions regarding AI interviewers, and may also partly explain why Study 2 did not observe significant within-person variance in privacy concerns (i.e., applicants' privacy concerns about AI interviewers remained constant and were not influenced by specific AI interviewers). Therefore, future research could extend the research sample to other countries and cultural groups, compare applicants' response patterns to AI interviewer image cues across different cultural contexts, and thereby further test the cross-cultural robustness of the findings.

Fourth, this study still has certain limitations in manipulating the image professionalism of AI interviewers. In Studies 1 and 3, the professionalism ratings of the low image-professionalism group on a 5-point scale all exceeded 4, resulting in a relatively small mean difference between the high and low groups. This indicates that the low image-professionalism group was not "completely unprofessional," but still possessed basic interviewer identity characteristics and a sense of professionalism. This treatment is consistent with the realistic requirements of recruitment contexts: as representatives of the recruiting party, AI interviewers, even under relatively low image-professionalism conditions, still need to maintain a basic professional image to ensure situational authenticity and participants' sense of immersion as applicants. The real AI-interview data in Study 2 also showed that participants' ratings of AI interviewer image professionalism were generally high ($M = 4.37$, within-person standard deviation = 0.48, between-person standard deviation = 0.33), indicating that AI interviewers in real recruitment contexts typically possess relatively high basic professionalism. If future studies aim to enlarge between-group differences and

Setting the low-professionalism condition as an image that was clearly inconsistent with the interview context may have caused participants to experience a strong sense of detachment from the situation, making it difficult to maintain the immersion and sense of presence required for a scenario experiment, while also weakening the realism of the situation, the effectiveness of the manipulation, and the external validity of the study. On the other hand, Study 1 manipulated image professionalism through differences between formal and casual attire, whereas Study 3 controlled for the fact that all interviewers wore formal business attire. However, as AI interviews are increasingly being applied in different industries such as IT, art and design, engineering, and fitness (Niuke; Moka, 2025), the expression of a professional image may not be limited to formal business dress. Therefore, future research could combine different positions and industries, using industry-appropriate attire, hairstyles, facial expressions, and voices to manipulate image professionalism, and further distinguish the independent effects of various cues.

Fifth, this study examined only gender as a cue of surface-level similarity. Similarity is generally divided into surface-level similarity and deep-level similarity: the former mainly refers to observable demographic characteristics such as gender, race, ethnicity, socioeconomic background, and age (Harrison et al., 1998), whereas the latter involves psychological characteristics such as personality, values, interests, attitudes, and beliefs (Engle & Lord, 1997). This study found that gender congruence, as an intuitive surface-level similarity cue, can increase applicants' trust in AI interviewers; however, this effect may also stem from perceptions of deep-level similarity activated by surface-level cues. Research on gender stereotypes has shown that men are typically perceived as more competent and dominant, whereas women are perceived as warmer and more sympathetic (Tay et al., 2014). This stereotype also extends to perceptions of AI (Ahn et al., 2022; Eagly & Wood, 2012). Therefore, male applicants may form judgments of competence similarity based on male AI interviewers, whereas female applicants may form judgments of warmth similarity based on female AI interviewers. As AI is gradually being endowed with deep-level attributes such as personality and value orientation (Serapio-García et al., 2025), future research could further examine whether the two congruent pairings—male applicant–male AI interviewer and female applicant–female AI interviewer—respectively elicit perceptions of competence similarity and warmth similarity, and investigate how deep-level matching between AI interviewers and applicants in values, attitudes, personality, and other aspects influences trust and subsequent reactions.

Acknowledgments: We thank research assistants Liu Xiaoying and Ma Yilin of Wenzhou-Kean University for their assistance in the preliminary work of this study.

References

Abbasi, Z., Billsberry, J., & Todres, M. (2024). Empirical studies of the "similarity leads to attraction" hypothesis in workplace interactions: A systematic review. *Management Review Quarterly, 74*(2), 661–709.

Aguinis, H., & Bradley, K. J. (2014). Best practice recommendations for designing and implementing experimental vignette methodology studies. *Organizational Research Methods, 17*(4), 351–371.

Ahn, J., Kim, J., & Sung, Y. (2022). The effect of gender stereotypes on artificial intelligence recommendations. *Journal of Business Research, 141*, 50–59.

Baek, T. H., & Kim, M. (2023). Is ChatGPT scary good? How user motivations affect creepiness and trust in generative artificial intelligence. *Telematics and Informatics, 83*, 102030.

Beal, D. J., & Weiss, H. M. (2003). Methods of ecological momentary assessment in organizational research. *Organizational Research Methods, 6*(4), 440–464.

Borau, S., Otterbring, T., Laporte, S., & Fosso Wamba, S. (2021). The most

human bot: Female gendering increases humanness perceptions of bots and acceptance of AI. *Psychology & Marketing, 38*(7), 1052–1068.

Brownlow, S. (1992). Seeing is believing: Facial appearance, credibility, and attitude change. *Journal of Nonverbal Behavior, 16*(2), 101–115.

Byrne, D. (1997). An overview (and underview) of research and theory within the attraction paradigm. *Journal of Social and Personal Relationships, 14*(3), 417–431.

Chen, Q. Q., & Park, H. J. (2021). How anthropomorphism affects trust in intelligent personal assistants. *Industrial Management & Data Systems, 121*(12), 2722–2737.

Chowdhury, S., Budhwar, P., Dey, P. K., Joel-Edgar, S., & Abadie, A. (2022). AI-employee collaboration and business performance: Integrating knowledge-based view, socio-technical systems and organisational socialisation framework. *Journal of Business Research, 144*, 31–49.

Contreras, I., Hossfeld, S., de Boer, K., Wiedler, J. T., & Ghidinelli, M. (2024). Revolutionising faculty development and continuing medical education through AI-Generated videos. *Journal of CME, 13*(1), 2434322.

da Motta Veiga, S. P., & Figueroa-Armijos, M. (2022). Considering artificial intelligence in hiring for cybervetting purposes. *Industrial and Organizational Psychology, 15*(3), 354–356.

da Motta Veiga, S. P., Figueroa-Armijos, M., & Clark, B. B. (2023). Seeming ethical makes you attractive: Unraveling how ethical perceptions of AI in hiring impacts organizational innovativeness and attractiveness. *Journal of Business Ethics, 186*(1), 199–216.

Eagly, A. H., & Wood, W. (2012). Social role theory. *Handbook of theories of social psychology, 2*, 458–476.

Engle, E. M., & Lord, R. G. (1997). Implicit theories, self-schemas, and leader-member exchange. *Academy of Management Journal, 40*(4), 988–1010.

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.

Feldkamp, T., Langer, M., Wies, L., & König, C. J. (2024). Justice, trust, and moral judgements when personnel selection is supported by algorithms. *European Journal of Work and Organizational Psychology, 33*(2), 130–145.

Foulk, T. A., & Lanaj, K. (2022). With great power comes more job demands: The dynamic effects of experienced power on perceived job demands and their discordant effects on employee outcomes. *Journal of Applied Psychology, 107*(2), 263–278.

Gambino, A., Fox, J., & Ratan, R. A. (2020). Building a stronger CASA: Extending the computers are social actors paradigm. *Human-Machine Communication, 1*, 71–85.

Gillespie, N., Lockey, S., Curtis, C., Pool, J., & Akbari, A. (2023). Trust in artificial intelligence: A global study. *The University of Queensland and KPMG Australia.*

Glikson, E., & Woolley, A. W. (2020). Human trust in artificial intelligence: Review of empirical research. *Academy of Management Annals, 14*(2), 627–660.

Goldberg, C. B. (2005). Relational demography and similarity-attraction in interview assessments and subsequent offer decisions: Are we missing something? *Group & Organization Management, 30*(6), 597–624

Graves, L. M., & Powell, G. N. (1995). The effect of sex similarity on recruiters' evaluations of actual applicants: a test of the similarity-attraction paradigm. *Personnel Psychology, 48*(1), 85–98.

Harrison, D. A., Price, K. H., & Bell, M. P. (1998). Beyond relational demography: Time and the effects of surface-and deep-level diversity on work group cohesion. *Academy of Management Journal, 41*(1), 96–107.

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.

Hunkenschroer, A. L., & Kriebitz, A. (2023). Is AI recruiting (un) ethical? A human rights perspective on the use of AI for hiring. *AI and Ethics, 3*(1), 199–213.

Hunkenschroer, A. L., & Lütge, C. (2022). Ethics of AI-enabled recruiting and selection: A review and research agenda. *Journal of Business Ethics, 178*(4), 977–1007.

Jin, E., & Eastin, M. (2024). Gender bias in virtual doctor interactions: Gender matching effects of chatbots and users on communication satisfactions and future intentions to use the chatbot. *International Journal of Human–Computer Interaction, 40*(23), 8246–8258.

Jones, A. C., Repke, A., Batastini, A. B., Sacco, D., Dahlen, E. R., & Mohn, R. S. (2023). The power of presentation: How attire, cosmetics, and posture impact the source credibility of women expert witnesses. *Journal of Forensic Sciences, 68*(3), 962–971.

Koch-Bayram, I. F., & Kaibel, C. (2024). Algorithms in personnel selection, applicants' attributions about organizations' intents and organizational attractiveness: An experimental study. *Human Resource Management Journal, 34*(3), 733–752.

Köchling, A., Wehner, M. C., & Warkocz, J. (2023). Can I show my skills? Affective responses to artificial intelligence in the recruitment process. *Review*

of Managerial Science, 17(6), 2109–2138.

Komiak, S. Y., & Benbasat, I. (2006). The effects of personalization and familiarity on trust and adoption of recommendation agents. *MIS Quarterly, 30*(4), 941–960.

Langer, M., & König, C. J. (2018). Introducing and testing the creepiness of situation scale (CRoSS). *Frontiers in Psychology, 9*, 2220.

Langer, M., König, C. J., Back, C., & Hemsing, V. (2023). Trust in artificial intelligence: Comparing trust processes between human and automated trustees in light of unfair bias. *Journal of Business and Psychology, 38*(3), 493–508.

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.

Lavanchy, M., Reichert, P., Narayanan, J., & Savani, K. (2023). Applicants' fairness perceptions of algorithm-driven hiring procedures. *Journal of Business Ethics, 188*(1), 125–150.

Lee, E. J., Nass, C. I., & Brave, S. (2000). Can computer-generated speech have gender? An experimental test of gender stereotype. In *CHI' 00 extended abstracts on Human factors in computing systems* (pp. 289–290).

Leo, X., & Huh, Y. E. (2020). Who gets the blame for service failures? Attribution of responsibility toward robot versus human service providers and service firms. *Computers in Human Behavior, 113*, 106520.

Li, M., & Suh, A. (2022). Anthropomorphism in AI-enabled technology: A literature review. *Electronic Markets, 32*(4), 2245–2275.

Lim, G. J., & Chua, R. Y. (2026). Through the lens of clarity: perceived organizational tightness boosts creativity for men, but not for women. *Organization Science, 37*(4), 1458–1486.

Liu, D., Zhang, Z., & Wang, M. (2018). Single level and multilevel moderated mediation and mediated moderation: Theoretical construction and modeling testing. In Chen, X. P., & Shen, W (Eds.), *Empirical Methods in Organization and Management Research* (3rd edition, pp. 663–694). Beijing: Peking University Press.
[刘东，张震，汪默. (2018). Single-level and multilevel moderated mediation and moderated mediation: theoretical construction and model testing (translated by Gao Zhonghua). In: Chen Xiaoping, Shen Wei (Eds.), *Empirical Methods in Organization and Management Research* (3rd edition, pp. 663–694). Beijing: Peking University Press.]

Lukacik, E. R., Bourdage, J. S., & Roulin, N. (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), 100789.

Maeda, T., & Quan-Haase, A. (2024). When human-AI interactions become parasocial: Agency and anthropomorphism in affective design. In *Proceedings of the 2024 ACM Conference on Fairness, Accountability, and Transparency* (pp. 1068–1077).

McCarthy, J. M., Bauer, T. N., Truxillo, D. M., Anderson, N. R., Costa, A. C., & Ahmed, S. M. (2017). Applicant perspectives during selection: A review addressing "So what?," "What' s new?," and "Where to next?" . *Journal of Management, 43*(6), 1693–1725.

Miles, S. J., & McCamey, R. (2018). The candidate experience: Is it damaging your employer brand?. *Business Horizons, 61*(5), 755–764.

Mirowska, A., & Arsenyan, J. (2025). The A (I) team: effects of human-likeness and conformity to gender stereotypes on initial trust and willingness to work with an AI teammate. *Journal of Organizational Behavior*, 1–25.

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.

Mori, M., MacDorman, K. F., & Kageki, N. (2012). The uncanny valley [from the field]. *IEEE Robotics & Automation Magazine, 19*(2), 98–100.

Mori, M., Sassetti, S., Cavaliere, V., & Bonti, M. (2025). A systematic literature review on artificial intelligence in recruiting and selection: a matter of ethics. *Personnel Review, 54*(3), 854–878.

Muthén, B. O., & Muthén, L. K. (2010). Mplus users guide. Sixth Edition Los Angeles. *CA: Muthén & Muthén*.

Nass, C., & Moon, Y. (2000). Machines and mindlessness: Social responses to computers. *Journal of Social Issues, 56*(1), 81–103.

Newell, S. J., & Goldsmith, R. E. (2001). The development of a scale to measure perceived corporate credibility. *Journal of Business Research, 52*(3), 235–247.

Nowcoder, & Moka. (2025). Campus recruitment data in autumn 2025. https://hr.nowcoder.com/article/14059
[Niuke & Moka. (2025). 2025 Autumn Campus Recruitment White Paper. See https://hr.nowcoder.com/article/14059]

Palmer, C. L., & Peterson, R. D. (2016). Halo effects and the attractiveness premium in perceptions of political expertise. *American Politics Research, 44*(2), 353–382.

Pitardi, V., Bartikowski, B., Osburg, V. S., & Yoganathan, V. (2023). Effects of gender congruity in human-robot service interactions: The moderating role of masculinity. *International Journal of Information Management, 70*, 102489.

Podsakoff, N. P., Spoelma, T. M., Chawla, N., & Gabriel, A. S. (2019). What predicts within-person variance in applied psychology constructs? An empirical

examination. *Journal of Applied Psychology, 104*(6), 727–754.

Preacher, K. J., & Selig, J. P. (2012). Advantages of Monte Carlo confidence intervals for indirect effects. *Communication Methods and Measures, 6*(2), 77–98.

Ridgeway, C. L., & Smith-Lovin, L. (1999). The gender system and interaction. *Annual Review of Sociology, 25*(1), 191–216.

Ryan, A. M., & Ployhart, R. E. (2000). Applicants' perceptions of selection procedures and decisions: A critical review and agenda for the future. *Journal of Management, 26*(3), 565–606.

Sari, Y. M., Hayu, R. S., & Salim, M. (2021). The effect of trustworthiness, attractiveness, expertise, and popularity of celebrity endorsement. *Jurnal Manajemen dan Kewirausahaan, 9*(2), 163–172.

Serapio-García, G., Safdari, M., Crepy, C., Sun, L., Fitz, S., Romero, P., ⋯ & Matarić, M. (2025). A psychometric framework for evaluating and shaping personality traits in large language models. *Nature Machine Intelligence, 7*, 1–15.

Siegel, M., Breazeal, C., & Norton, M. I. (2009). Persuasive robotics: The influence of robot gender on human behavior. In *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2563–2568.

Tay, B. T. C., Jung, Y., & Park, T. (2014). When stereotypes meet robots: The double-edge sword of robot gender and personality in human–robot interaction. *Computers in Human Behavior, 38*, 75–84.

Trougakos, J. P., Hideg, I., Cheng, B. H., & Beal, D. J. (2014). Lunch breaks unpacked: The role of autonomy as a moderator of recovery during lunch. *Academy of Management Journal, 57*(2), 405–421.

Uggerslev, K. L., Fassina, N. E., & Kraichy, D. (2012). Recruiting through the stages: A meta-analytic test of predictors of applicant attraction at different stages of the recruiting process. *Personnel Psychology, 65*(3), 597–660.

Vimalkumar, M., Sharma, S. K., Singh, J. B., & Dwivedi, Y. K. (2021). 'Okay google, what about my privacy?' : User' s privacy perceptions and acceptance of voice based digital assistants. *Computers in Human Behavior, 120*, 106763.

West, S. G., & Hepworth, J. T. (1991). Statistical issues in the study of temporal data: Daily experiences. *Journal of Personality, 59*(3), 609–662.

Wilhelmy, A., Kleinmann, M., König, C. J., Melchers, K. G., & Truxillo, D. M. (2016). How and why do interviewers try to make impressions on applicants? A qualitative study. *Journal of Applied Psychology, 101*(3), 313–332.

Wen, Z. L., Huang, B. B., & Tang, D. D. (2018). Preliminary work for modeling questionnaire data. *Journal of Psychological Science, 41*(1), 204–210.
[Wen Zhonglin, Huang Binbin, & Tang Dandan. (2018). Preliminary work for

modeling questionnaire data. *Journal of Psychological Science, 41*(01), 204–210.]

Xing, Y., He, W., Zhang, J. Z., & Cao, G. (2023). AI privacy opinions between US and Chinese people. *Journal of Computer Information Systems, 63*(3), 492–506.

Young, I. P., Place, A. W., Rinehart, J. S., Jury, J. C., & Baits, D. F. (1997). Teacher recruitment: A test of the similarity-attraction hypothesis for race and sex. *Educational Administration Quarterly, 33*(1), 86–106.

Zhang, L., & Yencha, C. (2022). Examining perceptions towards hiring algorithms. *Technology in Society, 68*, 101848.

Zhou, H., & Long, L. R. (2004). Statistical remedies for common method biases. *Advances in Psychological Science, 12*(06), 942–950.

[Zhou Hao, Long Lirong. (2004). Statistical testing and control methods for common method bias. *Advances in Psychological Science, 12*(06), 942–950.]

Zhou, L. L., Zhao, Y. S., & Zhao, S. M. (2025). The impact of anthropomorphism in artificial intelligence agents on human-agent trust and potential moderating effects: A meta-analysis. *Chinese Journal of Management, 22*(10), 1860–1871.

[Zhou Lulu, Zhao Yishan, Zhao Shuming. (2025). A meta-analytic study of the impact of anthropomorphism in artificial intelligence agents on human–agent trust and its potential moderating effects. *Chinese Journal of Management, 22*(10), 1860–1871.]

Appendix: Examples of the External Appearance of Corporate AI Interviewers

AI Interview System	External Appearance of the AI Interviewer	Companies Using It	Companies Using It
Hina AI	[Image: female AI interviewer in a beige suit; button text: “Complete interview and exit”]	OPPO, Meituan, etc.	Customizable dedicated AI interviewer appearance
Jinyu Intelligence	[Image: female AI interviewer portrait]	Siemens, JDB, etc.	

AI Interview System	External Appearance of the AI Interviewer	Companies Using It	Companies Using It
Niuke AI	[Image: female AI interviewer in a blue top]	CNOOC, counselor interviews, etc.	
AI Yimian	[Image: one female and one male AI interviewer]	China Post, China Minsheng Bank, etc.	Customizable dedicated AI interviewer appearance
Duomian AI	[Image: two AI interviewer interfaces; visible labels include “New male image” and “Classic female image”]	Life insurance companies, foreign-trade enterprises, etc.	
Yonyou Dayee	[Image: female AI interviewer in a white blazer]	Dongfeng Nissan, China Tower, Mengniu, etc.	Customizable dedicated AI interviewer appearance
Beisen AI	[Image: female AI interviewer wearing glasses; visible text includes “Interview has ended”]	Ruixing, etc.	

The Effects of AI Interviewers' Gender and Professional Appearance on

Job Applicant Reactions: The Mediating Role of Trust

WANG Lina[1], ZHENG Lu[1], Alexander Scott ENGLISH[2,3], ZHAO Teng[4], LI Hairong[5], JIANG Yan[1]

([1] School of Management, Huazhong University of Science and Technology, Wuhan 430074, China)

([2]Center for Behavioral Research Across Cultures in the School of Psychology, Wenzhou-Kean University, Wenzhou 325060, China)

([3]Sino-American Institute for Humanities and Economics, Wenzhou-Kean University, Wenzhou 325060, China)

([4]Business School, Nankai University, Tianjin 300071, China)

([5]School of Labor and Human Resources, Renmin University of China, Beijing 100872, China)

Abstract

AI interviews are becoming an increasingly integral component of organizational recruitment processes. Against this background, understanding how the appearance cues of AI interviewers shape applicant reactions is increasingly important. Drawing on similarity-attraction theory, this research examines whether gender congruence between AI interviewers and applicants enhances trust in the interviewer, thereby reducing applicants perceived creepiness of the interview process and privacy concerns. It further examines whether the AI interviewer' s professional appearance moderates these relationships. Specifically, when the interviewer appears less professional, applicants may rely more heavily on gender congruence to form trust judgments, thereby strengthening the indirect effects of gender congruence on perceived creepiness and privacy concerns through trust.

We conducted two scenario-based experiments and one experience sampling study to test the proposed hypotheses. Study 1 employed a 2(AI interviewer gender: male vs. female) $\times$ 2(professional appearance: high vs. low) between-subjects design. A pretest was conducted to select Chinese AI interviewer appearances for the formal experiment. Study 2 used an experience sampling design in which job-seeking students from five universities completed daily surveys over 30days to report their experiences with and perceptions of AI interviews. Study 3 replicated the 2×2 between-subjects design of Study 1 using non-Chinese AI interviewers selected through a pretest.

Study 1 showed that trust in the AI interviewer mediated the effects of gender congruence on both perceived creepiness of the interview process and privacy concerns. These indirect effects were further moderated by the interviewer' s professional appearance. When professional appearance was relatively low, applicants relied more heavily on gender-congruent cues, which strengthened the positive effect of gender

congruence on trust and, in turn, reduced perceived creepiness and privacy concerns. When professional appearance was relatively high, the effect of gender congruence was attenuated. Study 2 further showed that gender congruence enhanced trust in the interviewer, which subsequently reduced perceived creepiness. The findings again supported the moderating role of professional appearance. However, privacy concerns exhibited no significant within-person variance, suggesting that they remained relatively stable across AI interview experiences. Study 3 extended these findings to a cross-racial interview context: even when applicants interacted with non-Chinese AI interviewers, gender con-

gruence similarly fostered trust and, in turn, reduced perceived creepiness and privacy concerns. Across the three studies, professional appearance consistently moderated these relationships.

This research advances the literature on applicant reactions to AI interviews by shifting attention from functionality of AI interview to the appearance design of AI interviewers. It identifies gender congruence and professional appearance as two important appearance cues that shape applicants' trust and subsequent reactions. It also extends similarity-attraction theory to AI-mediated recruitment by showing that applicants engage in human-like social categorization processes when evaluating AI interviewers and that gender congruence serves as a similarity cue that strengthens trust. Finally, this research contributes to the literature on AI anthropomorphism by identifying a compensatory interaction mechanism among appearance cues. Specifically, professional appearance serves as a strong cue that anchors trust judgments and weakens the influence of gender congruence. When professional appearance is relatively low, applicants rely more heavily on gender similarity to reduce uncertainty.

Keywords: AI interviewer; Trust; Professional appearance; Applicant reactions; Similarity-attraction theory

Note: Figure translations are in progress. See original paper for figures.