

Between Consistency and Context: Public Perceptions of AI and Human Interview Evaluation

Ms. Suhana Aafreen Kareem Nawaz
Grade XII, DPS – Modern Indian School, Qatar

Abstract—Artificial intelligence is increasingly being considered for interview-based tasks like assessing and screening people. This study focuses on discovering how people perceive the features and functions of AI and humans in assessment procedures. A cross-sectional survey with 65 participants located in Qatar have completed a 26- question questionnaire on familiarity with AI, interview experience, as well as views on strengths and weaknesses, comfort, accountability, and opinions on AI vs. human evaluators. According to respondents, AI is correlated with consistency; factual evaluation; structured assessment; speed; and identifying patterns. Also, humans were more strongly associated with emotional understanding, contextual understanding, creativity, and personal interaction. Although respondents recognized possible benefits of AI (e.g., less personal bias and analytical help), concerns about privacy, algorithmic bias, incorrect evaluations, and limited contextual understanding remained common. Besides, mixed feelings about AI-assisted interviews were present, with discomfort more than comfort. In addition, most respondents preferred humans, shared responsibility, or situation-dependent responsibility for final evaluations rather than AI acting as the sole authority. Overall, the findings suggest that AI is viewed mainly as a supporting technology rather than a replacement for human judgment, highlighting the importance of human oversight and human-centered design. These results describe perceptions and associations in one cross-sectional sample; they do not establish that AI or humans perform better and do not demonstrate causal relationships.

Index Terms—Artificial intelligence; Human–AI interaction; Interview evaluation; Algorithmic decision-making; Public perception; Human judgment

I. INTRODUCTION

The prevalence of artificial intelligence is apparent in the human-centered assessment and communication phases, encompassing activities from the candidate

screening stage in the application process to actual interviewing processes that now make use of AI interviewers [4]. Such modern development allows some of the AI job interviewers to create questions automatically, provide information summaries, and revise candidates in a short period of time, which means that the time taken to accomplish all the activities involved in the interviews is greatly reduced [3][4]. Yet public trust in AI varies across tasks and contexts, with people frequently rating AI's capability higher than its benevolence [1]. Nonetheless, the level of distribution of the technology varies depending on the activities performed and requirements applied, [1]. However, these studies mostly evaluate AI performance or examine applicant and recruiter reactions, rather than asking how the general public perceives the division of capabilities between AI and humans [5].

The concept of artificial intelligence generates a considerable amount of disagreement. While AI can provide people with reliable and accurate information regarding objective matters, emotional and contextual aspects of human affairs associated to be addressed by human judgement [2][6]. AI assessment could be considered preferable for unambiguous and factual decisions but raising the questions about the distinctions between being able to "understand" a person and being able to "evaluate" one objectively adds more complexity to the topic of human monitoring [6]. In spite of the fact that artificial intelligence has been extensively studied in the context of evaluation, rather than asking how the general public perceives the division of capabilities between AI and humans [2].

Most of the previous research has concentrated on the performance of AI, the reactions of the people it evaluates, and the issues of organizational implementation of AI while the questions of

perception of shared responsibility remain under study [2][5]. The purpose of this research is to examine the attitudes of people towards the effectiveness of using AI in evaluation as opposed to that of human evaluator.

II. METHODOLOGY

This research employed a cross-sectional survey design to examine public perceptions of artificial intelligence (AI) and human capabilities in interview and evaluation processes. Data were collected using a structured, self-administered digital questionnaire.

A total of 65 respondents from the general public in Qatar participated in the survey. Participation was voluntary, and the questionnaire was distributed digitally.

Ethical Principles

The individuals involved in the research were able to give their approval to participate in the inquiry voluntarily as well as understood the aim of the research prior to being presupposed to fill the questionnaire. As for their participation, there was no need to disclose the name or any identification details, contact information, or any other information that could serve to identify them. The purpose of the research was shared with the participants beforehand, and they were entitled not to take part in the study if willing. All collected information was managed within the framework of the current study, in a collective manner and without revealing personal data about the participants. No financial support was attracted for the implementation of the study from outside sources. Respondents were categorized into the following age groups recorded in the survey:

- Below 15
- 15–18
- 18 or Above

There were no exclusion criteria aside from declining consent.

The questionnaire consisted of demographic, experience-based, perception-based, and opinion-based questions relating to AI-assisted interviewing.

The survey examined the following areas:

1. Respondent Profile and AI Familiarity

- Age group
- Familiarity with AI tools

2. Interview Experience

- Previous participation in interviews
- Awareness and use of AI for interview preparation

3. Perceptions of Interviews

- What interviews reveal about individuals?
- Factors affecting interview performance.
- Perceived consistency that technology could introduce.

4. AI Capabilities and Evaluation

- Qualities respondents would be comfortable or uncomfortable having technology evaluated.
- Expected capabilities of AI systems when analyzing interview responses.

- Expected similarity between AI and human evaluations.

5. Human–AI Capability Comparison

- Factual knowledge
- Consistency of evaluation criteria
- Understanding emotions
- Understanding context
- Identifying communication patterns
- Evaluating creativity

6. Benefits and Concerns

- Potential benefits of AI-assisted interviews
- Desired outcomes of technological assistance
- Concerns regarding AI use, including privacy, bias, human interaction, and inaccurate evaluation.

7. Human Oversight and Acceptance

- Comfort with AI involvement.
- Preferred responsibility for final evaluation
- Preferences when AI and human evaluations disagree.

8. Technology and Interview Improvement

- Areas in which technology could improve interviews.
- Suggestions for improving the modern interviewing process.
- Perceptions of whether understanding a person and evaluating them objectively are equivalent.

The complete questionnaire is provided in the Appendix.

Prior to analysis, data were cleaned using the following standardized procedures:

Data were analyzed using IBM SPSS All statistical tests were two-tailed with a significance level of $\alpha = 0.05$. The analytical approach comprised the following procedures:

1. Descriptive Statistics

- Frequencies (n) and percentages (%) were computed for every survey question, summarizing overall response patterns.
- For multiple-response questions, percentages were calculated based on the total number of respondents (N = 65) and may therefore exceed 100%.
- Descriptive statistics were used to summarize response distributions across all items.

2. Cross-Tabulation (Contingency Tables)

- Relationships between pairs of categorical variables — for example, AI familiarity × comfort with AI assistance — were examined using contingency tables.

3. Chi-Square Test of Independence (χ^2)

- Used to determine whether statistically significant associations existed between categorical variables.
- Because the sample was small (N = 65) and several questions had many response categories, Monte Carlo p-values were used alongside χ^2 whenever expected cell frequencies were too small for the ordinary chi-square approximation to be dependable.

4. Cramér's V

- Reported as the effect size for chi-square tests, indicating the strength of association independently of statistical significance.

5. Spearman's Rank-Order Correlation (ρ)

- Used for variables with a natural ordinal order — AI familiarity, comfort with AI assistance, and expected similarity between AI and human evaluations.
- Preferred over Pearson correlation because these variables are ordinal rather than truly continuous.

6. Kruskal–Wallis Test

- Used as a non-parametric comparison of an ordinal variable (e.g., comfort) across categorical groups, such as comparing comfort levels across AI-familiarity groups.

7. Qualitative Thematic Analysis

- Open-ended responses were analyzed thematically; recurring themes were identified and their frequencies reported where possible.

III. RESULTS

A total of 65 respondents completed the survey after giving consent. Participants were drawn from the general public in Doha, Qatar. The sample included respondents with varying levels of AI familiarity and interview experience.

A. Respondent Characteristics

A summary of demographic characteristics is shown below:

- Age distribution: Below 15 (~20%), 15–18 (~43%), Above 18 (~37%)
- AI familiarity: Very familiar ~52%, Moderately familiar ~37%, Slightly familiar ~9%, Not familiar ~2%
- Interview experience: ~55% had been in an interview at least once; ~21% had not; ~5% were unsure.

These demographic factors were used primarily for descriptive purposes.

B. Perceived Capabilities: AI vs. Human

Respondents associated AI and humans with clearly different strengths.

AI/Technology was most strongly associated with:

- Maintaining consistent evaluation criteria — 75.4%
- Evaluating factual knowledge — 67.7%
- Comparing responses using set criteria — 58.5%
- Identifying key information — 53.8%

Humans were most strongly associated with:

- Understanding emotions — 92.3%
- Evaluating creativity — 73.8%
- Understanding context — 63.1%
- Interpreting emotional responses — 76.9%

One area was divided: identifying communication patterns (AI 43.1% vs. human 41.5%).

22) Which would you expect to perform better at the following tasks?

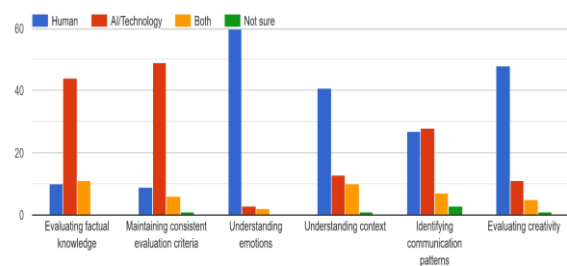


Fig. 1. Respondents' expectations of human versus AI/technology performance across interview tasks.

C. Comfort and Trust

Respondents showed a generally cautious attitude toward AI assistance:

- Comfortable with AI assistance: ~21.5% (somewhat + very)
- Uncomfortable: ~41.6% (somewhat + very)
- Neutral: ~23.1%
- Need more information: ~13.8%

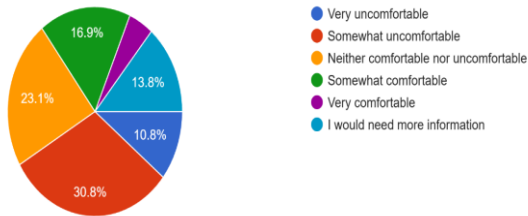


Fig. 2. Comfort levels with AI-assisted interview evaluation.

Main concerns: Lack of human interaction (33.8%), privacy of interview data (20.0%), algorithmic bias (15.4%).

9) Which quality would you be most comfortable having technology evaluate?
65 responses

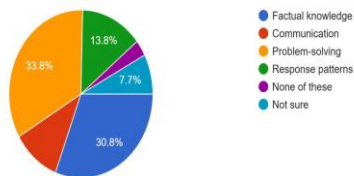


Fig. 3. Quality respondents would be most comfortable having technology evaluate.

D. Human Oversight and Responsibility

Respondents preferred human involvement and combined decision-making over autonomous AI:

- Human interviewer as final decision-maker: 40.0%
- Depending on situation: 26.2%
- Both equally: 21.5%
- AI system: 7.7%

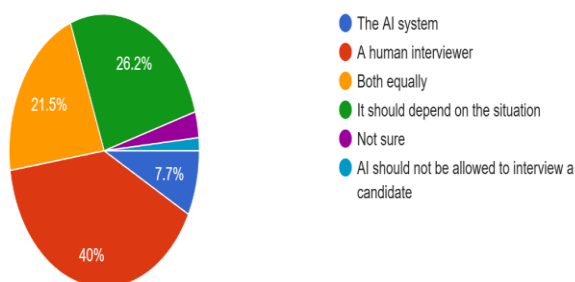


Fig. 4. Preferred responsibility for final evaluation when AI is used.

When AI and human evaluations disagree:

- Review both before deciding: 55.4%

Follow human judgment: 23.1%.

Follow AI: 10.8%

10) Which quality would you be least comfortable having technology evaluate?
65 responses

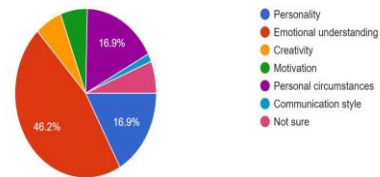


Fig. 5. Quality respondents would be least comfortable having technology evaluate.

E. Understanding vs. Objective Evaluation

75.4% of respondents said understanding a person is "No" or "Not necessarily" the same as evaluating them objectively which is a key conceptual finding.

F. Statistical Associations (Exploratory)

Because the sample is small (65) and several tables had sparse cells, associations were evaluated with Monte Carlo p-values and are treated as exploratory, not proven:

- AI familiarity $\times$ final responsibility: $\chi^2(15) = 72.444$, Monte Carlo $p = .0082$, $V = .610$ (strong, but sparse)
- AI preparation experience $\times$ comfort: $\chi^2(25) = 43.503$, $p = .0137$, $V = .366$
- Awareness $\times$ comfort: $\chi^2(15) = 25.971$, $p = .0356$, $V = .365$
- AI familiarity $\times$ expected AI-human similarity: Spearman $\rho = -0.290$, $p = .039$ — more AI-familiar respondents expected greater difference between AI and human evaluations

These are associations, not evidence that AI familiarity causes any attitude.

G. Qualitative Themes

Manual review of open-ended responses identified recurring themes:

Suggestions for improving interviews:

- Reducing interviewer bias/fairness (17 responses)
- Better question design and structure (12)
- Clearer role for technology/AI (13)
- Real-world skills over memorization (6)
- Keeping human judgment / human-AI balance (5)

These themes align with the quantitative findings.

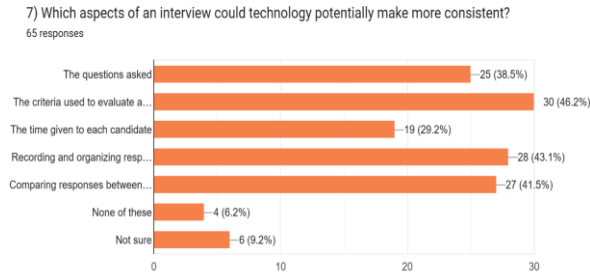


Fig. 6. Respondents' views on aspects of interviews that technology could make more consistent.

IV. DISCUSSION

This study investigated public perceptions about AI technology and human capabilities in the process of interviews and evaluations of candidates among respondents in Qatar. The results obtained align with the conclusions drawn in literature, that society treats AI and human judgment as having different, complementary strengths, rather than being interchangeable.

In the research sample, the division between AI and human abilities was evident. For instance, a majority of subjects associated AI/technology with consistent application of evaluation criteria (75.4% of respondents) and factual knowledge evaluation (67.7%). In contrast, human ability was viewed in light of emotional understanding (92.3%), creativity evaluation (73.8%), and context understanding (63.1%). All these trends at the respondents' level prove that technology is strong in structured and rule-based evaluations, while human judgment is believed to be more efficient in cases of interpretation which requires emotional and contextual understanding.

Strikingly, one specific domain was most contentious: identifying communication styles, whereby AI (43.1%) and humans (41.5%) are on a nearly equal footing. This infers that respondents do not perceive a straightforward separation of "AI versus human", but rather an arrangement wherein some tasks are considered combinative in nature. A guarded view towards the involvement of AI in interviews was evident, as only one fifth (21.5%) were comfortable using AI technology, while 41.6% did not make use of it and 13.8% expressed willingness to learn more about the process. The biggest concern was the lack of human connection (33.8%), followed by privacy of

interview data (20.0%) and the last place among fears was algorithmic bias (15.4%). These concerns reflect that general public worry about technology replacing personal interaction and taking control over personal information. Certain preference for human control became visible when the questions related to responsibility were asked. Human interviewers have been chosen as the preferred final decision maker by 40.0% of respondents and only 7.7% chose AI system. The same trend remained when it was asked what should be done in a case of disagreement between AI and human interviewers. In such case, 55.4% of respondents expressed opinion that all decisions should be reviewed prior to making a decision. This indicates that respondents generally want AI to function as an assisting tool rather than an independent final judge.

The disparity between understanding and objective assessment was also apparent. Three-quarters of participants (75.4%) argued that understanding someone does not equate to assessing them in an objective manner. This demonstrates that participants appreciate human judgment for those parts of the interview requiring the understanding of a person, but at the same time acknowledge the importance of technology for the purposes of obtaining consistency and fact-based information. Overall, the findings emphasize that the possibilities of using a complementary human-AI approach to interviewing, where technology serves the purpose of providing consistency and structure while humans maintain the responsibility for emotional connection, context, and creative thinking. Practical implications might be the application of AI for standardizing evaluation criteria and comparing answers but leaving human interviewers in charge of making the final decision and performing those segments of the interview that require empathy and context.

This research adds to a growing collection of studies about the public perception of AI technology in the process of interviews. Despite the limited number of sampled subjects, the findings are aligned with the discussions of the pros and cons of AI in companies where human-to-computer interaction is needed and also support the need for caution and use of the technology under supervision.

V. CONCLUSION

This research investigated how people in Qatar perceived AI and human abilities in the context of interviews and assessments. The research shows that people perceive AI and humans as having different abilities, with humans seen as more creative while AI is perceived as better at ensuring consistent, fact-based, and structured evaluations. Additionally, most respondents showed a rather restrained attitude toward AI involvement in the decision-making process, expressing a strong preference to keep humans in charge of making decisions alongside AI.

There were concerns about these systems taking away human interaction, respect for privacy, and bias. Moreover, most respondents acknowledged that knowing a person and objectively evaluating them is not similar.

In light of the increasing use of artificial intelligence in hiring and assessment, the results of this research emphasize the importance of creating AI-assisted interviews where human decision-making remains intact. Methods such as employing artificial intelligence for uniformity and organized evaluation while allowing humans to make the ultimate judgments should provide some means to combine efficiency with fairness.

The work contributes to limited findings about people's attitude toward utilization of AI-based interview methods in Middle Eastern countries and points out necessity of further investigation of responsible application of technology in systems of evaluation. It would be useful to conduct future research involving larger samples of more diverse population and objective evaluation of AI and human interviewers' performance in order to gain a better understanding of the phenomenon in Qatar and similar surrounding countries.

APPENDIX: SURVEY QUESTIONNAIRE

Consent statement: I have read the information above and understand the purpose of this study. I consent to participate voluntarily.

☐ Yes, I consent

☐ No, I do not consent (If "No," the survey ends here.)

1) Age group:

- ☐ Below 15

- ☐ 15–18

- ☐ Above 18

2) How familiar are you with artificial intelligence (AI) tools?

- ☐ Very familiar
- ☐ Moderately familiar
- ☐ Slightly familiar
- ☐ Not familiar at all

3) Have you ever participated in an interview?

- ☐ Yes, more than once
- ☐ Yes, once
- ☐ Yes
- ☐ No
- ☐ Not sure

4) In your opinion, what can an interview reveal most effectively? (Select all that apply)

- ☐ Subject knowledge
- ☐ Communication skills
- ☐ Personality
- ☐ Confidence
- ☐ Problem-solving ability
- ☐ Motivation
- ☐ It depends on the type of interview

5) Which statement best reflects your view of interviews?

- ☐ Interviews are a useful way of understanding people
- ☐ Interviews are useful but have significant limitations
- ☐ Interviews reveal only certain aspects of a person
- ☐ Interviews are generally unreliable for judging people
- ☐ It depends on how clearly the questions are asked
- ☐ I am unsure

6) Which factor do you think can most affect how a person performs in an interview?

- ☐ Nervousness
- ☐ Personality
- ☐ Type of questions asked
- ☐ Interviewer's approach
- ☐ Familiarity with the subject
- ☐ Preparation

- ☐ Other: _____
- 7) Which aspects of an interview could technology help make more consistent? (Select all that apply)
- ☐ The questions asked
 - ☐ The criteria used to evaluate answers
 - ☐ Recording and organizing responses
 - ☐ Comparing responses between candidates
 - ☐ The time given to each candidate
 - ☐ None of these
 - ☐ Not sure
- 8) If technology made these aspects more consistent, what do you think would happen?
- ☐ It could reduce some forms of interviewer bias
 - ☐ It could make evaluations more comparable
 - ☐ It could have both positive and negative effects
 - ☐ It could make interviews less flexible
 - ☐ Not sure
- 9) Which quality would you be most comfortable having technology evaluate?
- ☐ Factual knowledge
 - ☐ Response patterns
 - ☐ Communication
 - ☐ Problem-solving
 - ☐ None of these
 - ☐ Not sure
- 10) Which quality would you be least comfortable having technology evaluate?
- ☐ Personality
 - ☐ Communication style
 - ☐ Personal circumstances
 - ☐ Emotional understanding
 - ☐ Motivation
 - ☐ Creativity
 - ☐ Not sure
- 11) Before this survey, were you aware that AI-based interview tools exist?
- ☐ Yes
 - ☐ No
 - ☐ I had heard of it but knew little about it
- 12) Have you ever used an AI tool to prepare for an interview?
- ☐ Yes, regularly
 - ☐ Yes, a few times
 - ☐ Yes, once
 - ☐ No, but I would consider using one
 - ☐ No, and I would not consider using one
 - ☐ Not sure
- 13) If you have used or would consider using AI for interview preparation, which uses would be most helpful? (Select all that apply)
- ☐ Generating possible interview questions
 - ☐ Practicing responses
 - ☐ Receiving feedback on answers
 - ☐ Identifying weaknesses in responses
 - ☐ Improving communication
 - ☐ Learning about the company/role
 - ☐ I would not use AI for interview preparation
- 14) If an AI system analyzed an interview response, what do you think it could do? (Select all that apply)
- ☐ Identify key information in the response
 - ☐ Compare responses using set criteria
 - ☐ Identify communication patterns
 - ☐ Evaluate problem-solving responses
 - ☐ Evaluate emotional responses
 - ☐ None of these
 - ☐ Not sure
- 15) Which aspects of an interview do you think may be difficult for AI to evaluate? (Select all that apply)
- ☐ Emotional responses
 - ☐ Humour or sarcasm
 - ☐ Cultural context
 - ☐ Personal circumstances
 - ☐ Creativity
 - ☐ Motivation
 - ☐ None of these
 - ☐ Not sure
- 16) If an AI system and a human interviewer evaluated the same interview, how similar do you think their evaluations would be?
- ☐ Very similar
 - ☐ Somewhat similar
 - ☐ Neither similar nor different
 - ☐ Somewhat different
 - ☐ Very different
 - ☐ Impossible to predict
- 17) Which potential benefit of AI-assisted interviews appeals to you most?
- ☐ Reduced personal bias

- ☐ More consistent evaluation
- ☐ Faster evaluation
- ☐ Better analysis of responses
- ☐ More structured interviews
- ☐ Greater privacy
- ☐ Greater transparency
- ☐ None of these
- ☐ Not sure

18) If technology were used to assist an interview, which outcome would you most want?

- ☐ Reduced personal bias
- ☐ More consistent evaluation
- ☐ Faster evaluation
- ☐ Better analysis of responses
- ☐ More structured interviews
- ☐ Greater privacy
- ☐ Greater transparency

19) What would concern you most about the use of AI in interviews?

- ☐ Incorrect evaluation
- ☐ Algorithmic bias
- ☐ Privacy of interview data
- ☐ Lack of human interaction
- ☐ Over-reliance on technology
- ☐ Difficulty understanding context

20) If AI were used during an interview, who should make the final evaluation?

- ☐ A human interviewer
- ☐ The AI system
- ☐ Both equally
- ☐ It should depend on the situation
- ☐ AI should not be allowed to interview a candidate
- ☐ Not sure

21) How comfortable would you be with an AI system evaluating you in an interview?

- ☐ Very comfortable
- ☐ Somewhat comfortable
- ☐ Neither comfortable nor uncomfortable
- ☐ Somewhat uncomfortable
- ☐ Very uncomfortable
- ☐ I would need more information

22) Which would you expect to perform better at the following?

Maintaining consistent evaluation criteria:

- ☐ Human
- ☐ AI/Technology
- ☐ Both
- ☐ Not sure

Evaluating factual knowledge:

- ☐ Human
- ☐ AI/Technology
- ☐ Both
- ☐ Not sure

Understanding emotions:

- ☐ Human
- ☐ AI/Technology
- ☐ Both
- ☐ Not sure

Understanding context:

- ☐ Human
- ☐ AI/Technology
- ☐ Both
- ☐ Not sure

Identifying communication patterns:

- ☐ Human
- ☐ AI/Technology
- ☐ Both
- ☐ Not sure

Evaluating creativity:

- ☐ Human
- ☐ AI/Technology
- ☐ Both
- ☐ Not sure

23) If an AI system and a human interviewer disagreed, what should happen?

- ☐ Review both before deciding
- ☐ Follow the AI's evaluation
- ☐ Follow the human interviewer's judgment
- ☐ Conduct another evaluation
- ☐ Not sure

24) If an interview could be improved using technology, which improvement would be most valuable?

- ☐ Reducing interviewer bias
- ☐ Understanding responses better

- ☐ Improving evaluation accuracy
- ☐ Making questions more consistent
- ☐ Protecting candidate privacy
- ☐ Saving time
- ☐ I do not think technology would improve it

25) If you could change one aspect of the modern interviewing process, what would it be? (Open-ended)

26) In interviews, is understanding someone the same as evaluating them objectively?

- ☐ Yes
- ☐ No
- ☐ Not necessarily
- ☐ I am not sure

"Automation bias" and "selective adherence" to algorithmic advice. *Journal of Public Administration Research and Theory*.

- [9] Agudo, U., Liberal, K. G., Arrese, M., & Matute, H. (2024). The impact of AI errors in a human-in-the-loop process. *Cognitive Research: Principles and Implications*.
- [10] Vössing, M., Köhl, N., Lind, M., & Satzger, G. (2022). Designing transparency for effective human-AI collaboration. *Information Systems Frontiers*.

REFERENCES

- [1] Novozhilova, E., Mays, K. K., Paik, S., & Katz, J. E. (2024). More capable, less benevolent: Trust perceptions of AI systems across societal contexts. *Machine Learning and Knowledge Extraction*.
- [2] Lai, V., Chen, C., Liao, Q. V., Smith, A., & Tan, C. (2021). Towards a science of human-AI decision making: A survey of empirical studies. *arXiv (Cornell University)*.
- [3] Panfilova, A. S., et al. (2026). The AI interviewer: Multi-faceted evaluation of adaptive questioning by large language models. *Scientific Reports*.
- [4] Sri Latha, C., Krishna Kalyan, N., Abhilash, J., Kaushik, O. V. S. N., Balmiki, V., & Venukumar, S. (2023). Automated interview evaluation. *E3S Web of Conferences*.
- [5] Almeida, F., Junça Silva, A., Lopes, S. L., & Braz, I. (2025). Understanding recruiters' acceptance of artificial intelligence: Insights from the Technology Acceptance Model. *Applied Sciences*.
- [6] Binns, R. (2020). Human judgment in algorithmic loops: Individual justice and automated decision-making. *Regulation & Governance*.
- [7] Leybzon, D. D., Tirumala, S., Jain, N., Gillen, S., Jackson, M. R., McPhee, C., & Schmidt, J. E. (2025). AI telephone surveying: Automating quantitative data collection with an AI interviewer. *arXiv (Cornell University)*.
- [8] Alon-Barkat, S., & Busuioc, M. (2022). Human-AI interactions in public sector decision making: