Human-Computer Interaction & Emerging Technologies, Vol. 233, 2026, 81–89
https://doi.org/10.54941/ahfe1007506

The Impact of Direct and Third-Party Control: A Comparison of the Usage of AI Advice in Hiring Decisions

Johannes Zysk, Antonia Markus, Esther Borowski, and Ingrid Isenhardt

Chair for Intelligence in Quality Sensing (IQS), WZL | RWTH Aachen University, Aachen, Germany

ABSTRACT

For the trustworthiness of AI-based systems and their usage, control plays an important role from both regulatory and end-user perspectives. In general, two control approaches can be distinguished: direct control, giving the end user greater influence over the AI system, or a more indirect approach, by involving third parties to exercise control over the system. In this between-subjects experiment with 181 participants, four conditions with direct and indirect (third-party) control measures were compared in their usage of the AI systems' recommendations. During this study, participants evaluated the fit of fictional applications for a job opening. To assess system usage, we used the weight of advice (WOA), measuring the extent to which recommendations were considered in participants' assessments. A one-way ANOVA found a significant difference in the WOA between the levels of control: $F(3, 183) = 2.81$, $p = .041$. Group comparisons via contrasts showed a significant difference between the comparator and the third-party verification group (0.27; $SE = 0.10$; $p = .011$). Descriptively, all three experimental groups showed a higher usage (WOA) than the comparator group. This study shows the potential for control measures to deliver more trustworthy AI systems that see a higher usage of their recommendations. Thus, it provides practical implications for future design of AI-based decision support systems.

Keywords: Human-AI interaction, Human-AI collaboration, Trustworthy AI, Control

INTRODUCTION

While AI-based systems are becoming increasingly present in our day-to-day lives and work, their trustworthiness and usage play an essential role in effective human-AI collaboration. Control is a vital factor, to facilitate both the trust in AI systems and their usage.

One of the seven principles for trustworthy AI, as defined by the AI high-level expert group of the European Union, is human agency and oversight (AI HLEG, 2019). Here, the importance of control is highlighted, clearly stating the importance of user autonomy, as well as the necessity of oversight and control measures. This need is also mirrored on the user side. When asked about factors that would increase their trust in AI systems, and multiple answers were possible, two of the top three user entries concerned control (Salesforce, 2024). While the most frequently mentioned entry was on

Received February 6, 2026; Revised March 31, 2026; Accepted April 19, 2026; Available online July 20, 2026

transparency (42%), “Human validation of outputs” was the second-most frequently mentioned, at 35% of participants, and “Greater user control over AI usage” ranked third, with 32% of participants highlighting this aspect as important (Salesforce, 2024).

Thus, the present study aims to contribute to the ongoing discourse on control measures for trustworthy AI systems by exploring how different control measures influence the usage of recommendations by an AI-based decision support system. Specifically, we investigate whether direct and indirect (third-party) control measures succeed in enhancing the usage of such a system.

In the following sections, the current state of research and the hypotheses derived from it for this study will be presented. Afterwards, the experiment, its methods, and the results of this study will be described and discussed.

THEORETICAL BACKGROUND

Human-AI collaboration can be seen as a continuum, with fully manual work/no automation at one end and full automation at the other (Parasuraman et al., 2008). Between these two extremes, the actual collaboration occurs, which can be divided into human-dominant and AI-dominant interactions, depending on the level of agency or control (Sun et al., 2023). In the field of human-dominant interactions, AI-based decision support systems can be placed. These systems offer users information or advice but let them decide what action to take and how to react (Jao, 2010). To ensure a working Human-AI collaboration, usage and trust of the system, and to overcome adverse effects like AI aversion, where participants prefer human recommendations, even though an AI system delivers a more accurate result (Rahman et al., 2023; Burton et al., 2020), a factor often discussed is control (Dietvorst et al., 2018; Lee et al., 2019; Westphal et al., 2023). To evaluate the effects of the control measures implemented in this study, we examined system usage by measuring the weight of advice (WOA) or how much the recommendations were taken into account. Although trust is an important prerequisite for system usage, they are not the same, as users can trust a system and yet still not use it (Turel & Gefen, 2013). In this study, we investigated the effects of control on users’ behaviour, which is why we chose system usage as the focus.

In general, control can be distinguished between direct and third-party control. Bandura (1997), for example, describes personal, proxy, and collective agency, of which we will focus on the first two. While personal agency describes the direct control a person has over a situation, proxy agency describes control as the indirect exertion of agency over a situation by having others (third parties) that act on my behalf. Regarding AI systems, control is often primarily discussed in terms of direct control. Here, giving users already small amounts of control resulted in a more positive evaluation of the algorithm (Dietvorst et al., 2018). Additionally, studies found that giving users greater control in the interaction with the AI system, by allowing them to adjust the system’s recommendations, resulted in higher trust and compliance with the system and thus usage (Westphal et al., 2023).

Some applications of human-in-the-loop approaches, on the other hand, can be classified as third-party control, in which a human supervisor is part of the decision process, providing approval before the AI system creates its output (Samad, 2023). A study by Reis et al. (2024) found that advice from a human using AI didn't differ significantly from purely AI-based advice and performed significantly worse than purely human advice in terms of rated reliability, empathy, and the willingness to follow the advice. On the other hand, customer reactions were more positive when human control over the AI system was clearly evident (Haupt et al., 2024). Overall, previous studies on third-party control in the context of AI systems have shown mixed results regarding the beneficial effects of involving a human third party who can exert control over the AI system.

Additionally, many previous studies focused on the medical domain or on domain-agnostic tasks, while this study focuses on a task from the work context and compares both direct and indirect control measures. Thus, the study at hand addresses the following research question: Can different control measures (direct and third-party control) increase the usage of recommendations from AI systems in the context of a hiring decision?

METHODS

This chapter gives an overview of the methods of this current study. It presents information about the sample, experimental procedure, study design and operationalization.

Sample

Of the total sample of 194 participants who completed the questionnaire, seven had a relative speed index greater than 2, indicating potential inattentive or superficial responses (Leiner, 2019). The responses of these participants were checked, and five were excluded due to extremely low response times (2 to 5,5 minutes for the whole study) or missing data. Of the 189 remaining participants, eight needed to be excluded because they had coincidentally given the same value for their first rating as the AI recommendation later, resulting in an undefined value for the WOA in that trial and thus missing values for the total WOA for that participant.

The final sample consisted of 181 participants. Of this, 131 (72.4 %) were female, 46 (25.4 %) were male, three (1.7 %) identified as diverse, and one participant (0.6 %) chose not to answer. The average age was 29.65 years (SD = 11.46, range = 18–66 years). The level of education was relatively high: seven people had a secondary school diploma (3.9%), and 62 participants had a general university entrance qualification (34.2%). Twelve participants had completed vocational training (6.6%). 49 participants (27.1%) had a bachelor's degree, 35 participants (19.3%) had a master's degree/diploma, and 12 participants (6.4%) had a doctorate. Four participants (2.2%) indicated "other."

Procedure

The study was designed as an online vignette study on SoSci Survey (Leiner, 2024), using a Wizard-of-Oz approach. While the participants believed they were interacting with a real AI system, the presented recommendations were hard-coded. After receiving information about data protection, voluntary participation, and the study duration, the participants were introduced to the study context. As a team lead, they were tasked with identifying the most suitable applicant for a job vacancy. As a basis for this, they received a job profile with task descriptions and required qualifications. One after another, the participants received three applications for the position to evaluate their fit with the job profile. The participants received explanations and background information about the fictional AI system TalentMatchPro that they were about to interact with. Afterwards, they were randomly assigned to one of four conditions, each with a different control measure, as explained under the study design in the next chapter. In all groups, participants first assessed the applicants' fit for the job profile, then received the recommendation, and were asked to provide a final assessment that could differ from or match the first assessment. After evaluating all three applications, participants answered demographic questions and received their debriefing.

Thus, the experiment consisted of three phases: Firstly, participants provided consent and were introduced to the study context and background information on the AI system and its workings. Secondly, they received the job profile and an application, provided their initial assessment, received the AI system's recommendation, and then provided their final assessment. This second phase was repeated three times for three different applicants. The job profile, however, stayed the same. The third phase contained demographic questions, as well as the debriefing, where participants were informed of the study's aim and about the fact that they didn't interact with a real AI system or colleague, and for group two, that they weren't able to influence the system's responses when setting its priorities.

Study Design and Hypothesis

The study compared four conditions, one comparator group and three experimental groups that differed in the extent of control the participants had in their interactions with the fictional AI system. The first group acted as the comparator group with the lowest level of control (**low control group** = 46 participants). Here, the participants received a recommendation from the alleged AI system, as explained above, with no further specifications. The second group was able to change the system's priorities and weights (e.g., applicants' work experience, communication skills, etc.) as well as the decision, whether the system should include further information (e.g., applicants' photos or results from a previous aptitude test), giving the participants more control over the system's operations (**direct control group** = 44 participants). In the third group, instead of interacting directly with the alleged AI system, the participants were told that the provided recommendation came from a colleague who used the introduced AI system to assist with his response (**third-party AI usage group** = 47 participants). Group four had the same setup as group one but received the additional information that an experienced colleague had verified the validity of the systems' recommendations before they were provided to the participants (**third-party verification group** = 44 participants). Thus, groups three and four fell in the category of third-party control.

Based on the findings from the previous literature, this design described above was used to test the following hypotheses in this study:

H1: The four conditions of control differ significantly from each other

H1.1: Usage of AI recommendations in the comparator group is lower than when participants are given greater direct control (direct control group)

H1.2: Usage of AI recommendations in the comparator group is lower than when the AI system is used by a third-party to reach the recommendation (third-party AI usage group)

H1.3: Usage of AI recommendations in the comparator group is lower than when the recommendations are verified by a third-party (third-party verification group)

H1.4: The direct control group differs significantly from the two third-party groups

Operationalization

In this study, the weight of advice (WOA) is used to measure the usage of the AI systems' recommendations. It refers to the extent to which a person changes their initial assessment in response to a given recommendation, thereby reflecting the strength of that advice's influence. The WOA is calculated by dividing the difference between the final and initial estimates by the difference between the advice and the initial estimate.

$$\text{WOA} = \frac{\textit{final estimate} - \textit{initial estimate}}{\textit{advice} - \textit{initial estimate}}$$

A WOA value of 0 indicates that participants completely ignored the advice and that their final estimate was the same as their initial estimate. In contrast, a WOA of 1 indicates that subjects completely abandoned their original estimate and fully adopted the advice. At the participant level, a mean WOA value was calculated across the three evaluated applicants.

RESULTS

We conducted a one-way ANOVA to assess the effects of different levels of control on the usage of the AI systems' recommendations (as measured by the WOA). A significance level of $\alpha = .05$ was set for the analyses. For the levels of control, there were four groups: direct control ($M = 0.30$, $SD = 0.50$), low control ($M = 0.06$, $SD = 0.64$), third-party AI usage ($M = 0.25$, $SD = 0.34$), and third-party verification ($M = 0.33$, $SD = 0.47$). Although the Shapiro-Wilk test indicated a violation of the normality assumption ($p < .05$), an ANOVA was conducted due to its inherent robustness to non-normal distributions. There was homogeneity of variance (Levene's test, $p > .05$). The usage of the AI systems' recommendations (as measured by the WOA) differed statistically significantly for the different levels of control: $F(3, 183) = 2.81$, $p = .041$, $\eta_p^2 = .044$.

Group Comparison Via Contrasts

To test the formulated subhypotheses H1.1, H1.2, and H1.3, contrasts were calculated comparing group 1 (low control) with groups 2 (direct control), 3 (third-party AI usage), and 4 (third-party verification). For the subhypothesis H1.4, an additional contrast was calculated to compare group 2 to the two third-party groups (groups 3 and 4). A Bonferroni-corrected p-value of p = .0125 was used for the calculation. The calculated contrasts are displayed in Table 1.

Table 1: Calculated contrasts.

	Contrast Values	Standard Error	p Value
Low control vs direct control	0.24	0.10	.023
Low control vs third-party AI usage	0.19	0.10	.068
Low control vs third-party verification	0.27	0.10	**.011***
Direct control vs third-party groups	0.01	0.09	.904

*statistically significant against a Bonferroni-corrected p-value of p = .0125

There was a statistically significant difference in WOA-scores between the groups with low control (M = 0.06, SD = 0.64) and third-party verification (M = 0.33, SD = 0.47) of 0.27 (SE = 0.10), p = .011. While the group with low control descriptively had the lowest value of all four groups (0.06 vs. 0.30; 0.25; 0.33), the other two contrasts comparing this group to groups 2 and 3 were not significant. Additionally, the contrast between the direct control group and the two third-party groups (3 and 4) found no significant difference.

According to Yaniv (2004), all WOA-values can be described as low, with values from 0 to 0.3 classified as low, values from 0.4 to 0.6 as medium, and values from 0.7 to 1.0 as high.

DISCUSSION

This study examined the effect of different control types on the usage of AI advice. For this, four groups were compared in a between-subjects study design. The conditions were a comparator group with low control and three experimental groups with different control measures. The first experimental group was a direct control group, in which participants had higher control because they could select priorities for the AI system and the other two groups were designed as third-party controls: in one condition, a colleague gave the assessment while using the AI system, and in the other, participants were notified that a colleague had previously verified the validity of the AI system's recommendations. In this study, hypotheses H1 (The four conditions of control differ significantly from each other) and H1.3 (Usage of recommendations in the comparator group is lower than in the third-party verification group) were supported by the data, finding statistically significant results. In contrast, hypotheses H1.1 (Usage of recommendations in the comparator group is lower than in the direct control group), H1.2 (Usage of recommendations in the comparator group is lower than in the third-party AI usage group), and H1.4 (The direct control group differs significantly from the two third-party groups) were not supported by this study's findings.

This study's findings contribute to the growing literature on human AI collaboration, showing that control measures are an important way to increase the use of recommendations from an AI-based decision support system. Its findings are thus in line with previous studies (Dietvorst et al., 2018; Westphal et al., 2023; Haupt et al., 2024). It particularly highlights the potential of third-party verification of AI results (H1.3). This aligns with previous studies that found end users seeking third-party verification and independent certificates to enhance their trust in an AI system (Werz et al., 2023). Descriptively, the comparator group with low control had the lowest usage of all four groups. As all WOA values were below 0.39, they can be rated as low (Yanniv, 2004). Here, previous studies have shown that algorithmic recommendations are taken into account much more strongly when presented before participants make a decision than after (Aiyer & Yeung, 2025; Serra-Garcia & Gneezy, 2024). However, this was necessary to calculate the WOA and thus to measure the system's usage.

Some results of this study however, were not expected from previous literature: Regarding the comparison of the comparator to the direct control group, it is possible that the participants felt like they already had a lot of control in the comparator condition, as the system was always a decision support system, falling in the human-dominant part of the human-AI collaboration continuum (Sun et al., 2023). In fact, other studies, such as that by Westphal et al. (2023), used a condition similar to the comparator in this study as their high-control condition. Additionally, the priority selection did not affect the recommendations during the study, which might have reduced the positive effect.

As we couldn't find a significant difference between the comparator and the third-party AI usage group, it may be that it wasn't clear that the alleged colleague was in control of the system, or that the participants didn't trust him to use that control as they wished. This might also explain the mixed results found in previous studies (Reis et al., 2024; Haupt et al., 2024). In this case, it might be crucial that it is very clear that the third party is in control of the system, how exactly they use it, and for what purpose. Also, in this case, the third party may not have been trusted enough to improve the AI system's perception.

Some additional limitations of this study should be noted. During the participants' trial, they only had three interactions with the system. This might not have given them enough time and opportunity to get a feel for the system and build trust in it. Additionally, the participants did not interact with a real AI system; instead, they received scripted answers in this vignette study. Consequently, the preferences that could be manipulated in one condition didn't change the given recommendations.

For future work, the areas of human-AI collaboration and the effects of control remain relevant and promising research fields. As shown in this study, third-party verification was an effective tool for enhancing system usage. As this might not always be possible in practice at the output level, it would be interesting to compare other types of verification and their effects on AI system perception and usage. This could entail, e.g., verification or certification of the system's overall reliability, automatic output verification by a third-party service, or on-demand verification of individual results. Despite not finding an effect in the condition with higher direct control, it might still be valuable to compare different options of enhanced direct control, other than the one chosen here, in which participants could manipulate system priorities at the start of the interaction. One possibility would be a different Human-in-the-loop approach, in which users provide feedback during the interaction, thereby shaping the system's answers. Overall, we recommend that future studies use longer or more frequent interactions with the AI system, and it might also be interesting to have participants interact with an actual AI system to enhance the study's realism.

CONCLUSION

This work investigated the effect of different control measures on the usage of an AI-based decision support system. Here, we found a statistically significant difference across the control conditions in the usage of the AI systems' recommendations. Contrast comparisons showed a significant difference between the comparator group and the third-party verification group. This shows the benefit of providing a verification of AI systems and their outputs by third parties, for the usage of AI recommendations. This study, with its results, contributes to the developing field of research on human-AI collaboration and carries practical implications for the design of trustworthy and usable AI.

ACKNOWLEDGMENT

Funded by the European Union. This work has received funding from the European High Performance Computing Joint Undertaking (JU) and from the German Federal Ministry of Research, Technology, and Space (BMFTR), the Ministry of Culture and Science of North Rhine-Westphalia (MKW NRW), and the Hessian Ministry of Science and Research, Arts and Culture (HMWK) under grant agreement No 101250682.

REFERENCES

AI HLEG. (2019). Ethics guidelines for trustworthy AI. European Commission. https://ec.europa.eu/digital-single-market/en/news/ethics-guidelines-trustworthy-ai

Aiyer, S., & Yeung, N. (2025). The impact of the timing of advice on its utilization. Journal of Behavioral Decision Making, 38:e70021. https://doi.org/10.1002/bdm.70021

Bandura, A. (1997). Self-efficacy: The exercise of control. W H Freeman/Times Books/ Henry Holt & Co.

Burton, J. W., Stein, M. K., & Jensen, T. B. (2020). A systematic review of algorithm aversion in augmented decision making. Journal of Behavioral Decision Making, 33(2), 220–239. http://dx.doi.org/10.1002/bdm.2155.

Dietvorst, B. J., Simmons, J. P., & Massey, C. (2018). Overcoming algorithm aversion: People will use imperfect algorithms if they can (even slightly) modify them. Management Science, 64(3), 1155–1170. http://dx.doi.org/10.1287/mnsc.2016.2643.

Haupt, M., Freidank, J., & Haas, A. (2024). Consumer responses to human-AI collaboration at organizational frontlines: strategies to escape algorithm aversion in content creation. Review of Managerial Science, 1–37. https://doi.org/10.1007/s11846-024-00748-y

Jao, C. (Ed.). (2010). Decision Support Systems. InTech. https://doi.org/10.5772/3453

Lee, M. K., Jain, A., Cha, H. J. I. N., Ojha, S., & Kusbit, D. (2019). Procedural justice in algorithmic fairness: Leveraging transparency and outcome control for fair algorithmic mediation. Proceedings of the ACM on Human-Computer Interaction, 3:182. https://doi.org/10.1145/3359284

Leiner, D. J. (2019). Too fast, too straight, too weird: Non-reactive indicators for meaningless data in internet surveys. Survey Research Methods, 13(3), 229–248. https://doi.org/10.18148/srm/2019.v13i3.7403

Parasuraman, R., Sheridan, T. B., & Wickens, C. D. (2008). Situation awareness, mental workload, and trust in automation: Viable, empirically supported cognitive engineering constructs. Journal of Cognitive Engineering and Decision Making, 2(2), 140–160. http://dx.doi.org/10.1518/155534308X284417.

Rahman, M. J., Liang, H., & Xue, Y. (2023). AI Aversion: A Task Dependent Multigroup Analysis. PACIS 2023 Proceedings. 86.

Reis, M., Reis, F., & Kunde, W. (2024). Influence of believed AI involvement on the perception of digital medical advice. Nature Medicine, 1–3. https://doi.org/10.1038/s41591-024-03180-7

Salesforce (2024). State of the AI Connected Customer. Seventh Edition. https://www.salesforce.com/en-us/wp-content/uploads/sites/4/documents/research/State-of-the-Connected-Customer.pdf

Samad, T. (2023). Human-in-the-loop control and cyber–physical–human systems: applications and categorization. Cyber–physical–human systems: fundamentals and applications, 1-23. https://doi.org/10.1002/9781119857433.ch1

Serra-Garcia, M., & Gneezy, U. (2024). Improving human deception detection using algorithmic feedback (10518). https://doi.org/10.1287/mnsc.2023.02792

Sun, Y., Shen, X.-L. & Zhang, K.Z.K. (2023). Human-AI interaction, Data and Information Management, Volume 7, Issue 3,100048, https://doi.org/10.1016/j.dim.2023.100048.

Turel, O., & Gefen, D. (2013). The dual role of trust in system use. Journal of Computer Information Systems, 54(1), 2-10. https://doi.org/10.1080/08874417.2013.11645666

Werz, J., Borowski, E., Isenhardt, I. (2024). Explainability as a means for transparency? Lay users' requirements towards transparent AI. In: Lucas Paletta (eds) Cognitive Computing and Internet of Things. AHFE (2024) International Conference. AHFE Open Access, vol 124. https ://doi.org/10.54941/ahfe1004712

Westphal, M., Vössing, M., Satzger, G., Yom-Tov, G. B., & Rafaeli, A. (2023). Decision control and explanations in human-AI collaboration: Improving user perceptions and compliance. Computers in Human Behavior, 144, 107714. https://doi.org/10.1016/j.chb.2023.107714

Yaniv, I. (2004). The benefit of additional opinions. Current Directions in Psychological Science, 13(2), 75–78. https ://doi.org/10.1111/j.0963-7214.2004.00278.x