



Recruiters' Behaviors Faced with Dual (AI and human) Recommendations in Personnel Selection

Alain Lacroux, Christelle Martin Lacroux

► To cite this version:

Alain Lacroux, Christelle Martin Lacroux. Recruiters' Behaviors Faced with Dual (AI and human) Recommendations in Personnel Selection. Academy of Management Annual Meeting, Academy of Management, Aug 2023, Boston (MA), United States. hal-04200429

HAL Id: hal-04200429

<https://paris1.hal.science/hal-04200429v1>

Submitted on 8 Sep 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

Recruiters' Behaviors Faced with Dual (AI and human) Recommendations in Personnel Selection**ABSTRACT**

Artificial Intelligence (AI) is increasingly used for decision-making support in organizations, and especially during the recruitment process. Consequently, recruiters may sometimes find themselves having to process different sources of information (human vs. algorithmic decision support system, ADSS) before deciding to preselect an applicant. Our study aims to explore the mechanisms that lead recruiters to follow or not the recommendations made by human and non-human experts, in particular when they receive contradictory or inaccurate information from these sources. Drawing on results obtained in the field of automated decision support, we make a first general hypothesis that recruiters trust human experts more than ADSS and rely more on their recommendations. Secondly, based on the Judge Advisor System Paradigm (Sniezek & Buckley, 1995), we make a second general hypothesis that the accuracy of the recommendations provided by the dual source of advice influences in different ways the accuracy of recruiters' preselection decisions. We conducted an experiment involving the screening of resumes by a sample of professionals (N=746) responsible for screening job applications in their work. As hypothesized, the recommendations made to recruiters do influence the accuracy of their decisions. Our results suggest that recruiters comply more with ADSS than human recommendations even if they declare a higher level of trust in human experts. Finally, implications for research and HR policies are discussed.

RECRUITERS' BEHAVIORS FACED WITH DUAL (AI AND HUMAN)

RECOMMENDATIONS IN PERSONNEL SELECTION

Artificial Intelligence (hereafter, AI) is nowadays present in all aspects of our lives, and especially our work life (De Cremer, 2020). Previous research has repeatedly highlighted the superiority of AI performance across multiple tasks and domains, and has demonstrated its potential benefits in terms of productivity and economic gains (Brynjolfsson & Mitchell, 2017; Grove et al., 2000). Recent advances also allow organizations to increasingly use AI-based solutions for complex tasks, including managerial functions such as decision making (Haesevoets et al., 2021).

Most organizational scholars consider AI to be more of a support mechanism than a substitute for complex managerial tasks, and are therefore interested in human-AI cooperation (Grønsund & Aanestad, 2020; Sowa et al., 2021), for instance in the selection process for recruitment interviews. Although it is generally acknowledged that most occupations will not be totally replaced by AI, many work environments will nevertheless experience some degree of reengineering and reorganization of work practices and tasks, with human workers and AI working side by side. This is likely to lead researchers to focus in the future on the study of the interdependence between humans and computers (Brynjolfsson & Mitchell, 2017; Solberg et al., 2022).

In the field of recruitment, AI is used at all stages of the process, from relatively simple tasks such as extraction of information from resumes, to highly complex and subjective tasks, such as the automated analysis of interviews or multi-criteria selection of the “best candidate” (Nawaz, 2020). Delegating the entire recruitment process to AI is not a realistic option; nor is the full automation of selection processes the goal of organizations. Instead, organizations and developers of predictive recruitment solutions prefer to use the term “augmented recruiting”,

defined as a process in which humans work closely with machines to complete a task (Raisch & Krakowski, 2021). The study of the mechanisms that lead recruiters to collaborate or not with AI embedded in Algorithmic Decision Support Systems (hereafter ADSS) is crucial in that it raises technical, managerial, ethical, and legal issues.

In this context of AI-recruiter collaboration, new questions emerge (König & Langer, 2022), particularly regarding the understanding of the perceptions and behavior of recruiters operating in an environment saturated with information and advice from both peers (other recruiters) and systems incorporating AI (chatbot, automatic recommendations of written applications, automatic analyses of asynchronous video interviews). The coexistence of different sources of information is a challenge for recruiters, who need to determine who to trust, and which information (human vs. algorithmic) to rely on to make a final decision, such as whether or not to invite a candidate to interview after the application assessment phase. One of the key issues that arises is when the recruiter receives contradictory information; in this case, the question is which source most influences their attitudes and behavior?

Given the extensive use of AI-based tools, this paper aims to extend previous research on the mechanisms that lead individuals to collaborate or not with AI when faced with different - possibly divergent - sources of recommendations in the selection process. It is now well established that during the selection process, trust plays a central role in recruiters' decisions when they receive recommendations (Chamorro-Premuzic et al., 2016; Gonzalez et al., 2019; Kuncel et al., 2013; Langer et al., 2019). Empirical research conducted on decision maker's trust in the context of a collaboration between humans and robots remains scarce in the field of Human Resource Management, with a gap in the knowledge on how recruiters perceive human and non-human recommendations, and how this perception influences their recruitment choices.

To contribute to filling this gap, our paper focuses on recruiters' perceptions and behaviors during the preselection of candidates, a particular stage in the recruitment process where AI - human collaboration is a growing phenomenon. In doing so, we look to address Glikson & Woolley's (2020) call for further research in AI-human collaboration in organizations adopting a "human centered approach".

We first present the conceptual basis for the study of trust in human experts and ADSS in judgmental situations such as recruitment, and develop our hypothesis on the attitude and behaviors of recruiters receiving dual sources of advice. We then analyze and discuss the results drawn from an experimental study conducted on a sample of recruiters (N = 788).

TRUST IN ADSS OR/AND HUMAN EXPERTS: THE CASE OF RECRUITMENT

Decisions made during the recruitment process can be viewed as a typical example of complex and partly subjective decisions processed by an expert. This expert is sometimes helped by human individuals (e.g., HR experts, co-workers, consultants) or technological devices (such as ADSS). In these situations, trust is considered to be a crucial predictor of both technology use (Dzindolet et al., 2002; Hoff & Bashir, 2015; Mayer et al., 1995) and the consideration of human sources of advice (Snizek & Buckley, 1995). Following Solberg et al. (2022), who elaborated a conceptual model of trust in AI-generated decision aids grounded in the model of trust in organizations developed by Mayer et al. (1995), we define trust as "the willingness to be vulnerable to the actions of another party, which is informed by the trustor's perceptions of the other party's trustworthiness and of their own general propensity to trust" (ibid, p 191). This emphasis on vulnerability as a core component of trust is particularly salient in the case of "augmented" personnel selection because it relates to a "black box problem" involving the inability for the trustor to decipher and monitor the algorithmic process that provides decision-making support. Vulnerability depends on the risk the trustor perceives in the trustee (the entity

trusted by the trustor), and therefore may differ depending on the nature of the trustee (human or machine).

Differences between algorithmic and human trust

Madhavan & Wiegmann (2007) found that trust in human-human interaction has some similarities with trust in human-automation interaction, but significant differences remain between human-ADSS trust and human-human trust. According to the above authors, people commonly assume that machines are perfect and perceive them as being able to perform consistently across situations. Consequently, trust in ADSS (a type of algorithmic machine) is based on a performance evaluation with a situational model of "perfection". Consequently, user trust is very sensitive to errors made by ADSS, and trust in ADSS rapidly dissolves when users are exposed to errors.

On the one hand, automated systems may trigger a positive machine heuristic, defined as a general belief that machines, compared to humans, are more reliable (M. K. Lee, 2018). This results in AI having greater influence on users, and reveals a preference for algorithmic judgement (Logg et al., 2019; You et al., 2022). Concerning the recommendation of business profiles, Wang et al. (2020) found that AI systems were just as influential as human experts. Other research has distinguished between the type of task undertaken by AI: for simple perceptual tasks, users report a high level of trust in AI, which leads them to follow erroneous recommendations (Liel & Zalmanson, 2020). People also exhibit greater algorithmic appreciation as intellectual tasks becomes more difficult. (Bogert et al., 2021).

On the other hand, Bogert et al. found in line with Madhavan and Wiegmann (2007) that when the quality of advice decreases, algorithms are penalized to a greater degree than a crowd of peers. This lack of trust may prevent the integration of AI systems and agents into teams (Groom & Nass, 2007). Such algorithmic aversion involves individuals not trusting machines and

therefore not wanting to rely on them (Dietvorst et al., 2015). One reason for this aversion lies in the fact that they do not understand how machines make decisions and cannot exert control over the process, meaning that many people still perceive AI as a “black box” (Gillath et al., 2021), and thus creating a challenge for trust in AI embedded decision aid (Solberg et al. 2022).

Concerning human experts, Madhavan and Weigmann (2007) provided evidence that they are perceived as being relatively more adaptable and having the ability to modify their behavior according to situations. Consequently, trust in human recommendation is based on an evaluation of knowledge with a dispositional model of "imperfection". The level of expertise of the human trustee is a major determinant of initial level of trust as it may induce an “authority heuristic”, which is proportional to the credibility of the source (Wang et al., 2020). In the case of human-human interaction, the user’s trust is less affected by mistakes made because human experts are expected to occasionally commit mistakes.

The impact of dual recommendations (ADSS & Human)

When it comes to the case of human-machine cooperation in managerial decision making, Haesevoets et al. (2021) showed that people prefer situations where the human agent has a dominant input (70%). In a context of mixed recommendations made by ADSS and human experts, Wang and colleagues (2020) pointed to the lack of consistency in empirical evidence comparing the influence of human experts versus ADSS on the judgement of decision makers.

When the decision made relies on advice received from other sources, scholars regularly refer to the “Judge Advisor System Paradigm” (Bailey et al., 2022; Pescetelli & Yeung, 2021; Snizek & Buckley, 1995) to analyze the way persons providing recommendations and information (the advisors) influence the person with the responsibility for the decision (the judge). Following an emerging trend of research studying “hybrid” judge advisor systems (Hou & Jung, 2021; Mesbah et al., 2021; Schemmer et al., 2022; You et al., 2022) , we argue that

this paradigm may be applied to situations where some advisors are non-human. This is the case with “AI-augmented” recruitment (Priksht et al., 2023), in which humans work closely with ADSS to complete a task (Raisch & Krakowski, 2021) and receive different sources of information from both humans and ADSS (Grønsund & Aanestad, 2020). Research conducted on judge advisor systems following the seminal work of Sniezek and Buckley (1989, 1995) has led to several key findings of particular interest for augmented recruitment. Firstly, even experienced judges accept advice, and the more experienced the advisor is considered to be, the more the judge will be influenced (Harvey & Fischer, 1997; Sniezek & Van Swol, 2001). However, judges tend to give too much weight to their own opinion in relation to that of their advisor (Harvey & Fischer, 1997; Yaniv, 2004). Secondly, information about the advisor that is suggestive of quality advice is the major predictor of advisor influence (Bailey et al., 2022). Thirdly, trust in the advisor is a central variable when the decision becomes more complex or uncertain (Sniezek & Van Swol, 2001). Consequently, advice-taking is greater for subjective and uncertain estimates (Bailey et al., 2022). Finally, in case of disagreement between advisors, the judge faces a cognitive conflict that may affect their confidence and accuracy. On the one hand, conflict is expected to benefit performance (all things being equal) because it encourages the judge to examine more closely the different alternatives (Sniezek et al., 1990). On the other hand, cognitive conflict may also reduce the judge’s confidence in their own choice. It should be noted that research on cognitive conflict is inconclusive because cognitive conflict has been shown to *reduce* accuracy of choice in Sniezek & Buckley (1995).

Trust and advice-taking in the context of “augmented” recruitment

Concerning the selection process, studies agree that recruiters demonstrate aversion by tending to distrust and avoid the use of automation. Rieger et al. (2022) revealed in various contexts that human recommendation is considered to be more trustworthy than AI recommendation, especially in cases of complex decision-making situations that are comparable to the selection

of personnel. Langer & Landers, (2021) also noticed that trustworthiness assessment could be lower for ADSS in tasks that may affect the fate of individuals such as in personnel selection. According to Langer, König and Busch (2021), recruiters perceive a human trustee as more trustworthy than an automated one. This preference for human recommendations emerges even if the recommendation is highly subjective (Oberst et al., 2020). If the general conclusion is that recruiters tend to place more trust in human experts, there remains uncertainty concerning the influence of ADSS on recruiters' behavior. Most of the previous studies used inter-subject designs, in which participants were exposed to only one type of recommendation (Oberst et al., 2020).

Our study aims to go further by investigating the influence of dual recommendations (human and algorithmic) on recruiters' preselection decisions by comparing the influence of each source of recommendation and analyzing the case of non-congruent or conflictual recommendations. Such situations are likely to become more commonplace in the near future because the rapid development of AI is likely to favor human-AI collaboration in organizational decision making tasks, leading to hybrid decision making processes (Shrestha et al., 2019).

Hypotheses

The present study aims to firstly test whether recruiters' trust and behavior are significantly influenced by the nature and accuracy of the advisor (human expert *versus* ADSS) in situations of dual sources of recommendation. We then focus on the particular case of disagreement between advisors, by asking which category of advisor is the most influential in such situations.

Hypotheses related to recruiters' preselection decisions:

In line with Harvey & Fisher (1997), we posit that recruiters tend to accept recommendations provided by advisors who are considered to be experts (whether human or non-human) and give

more weight to their judgement compared with that of advisors. We therefore develop the following hypotheses for situations where advisors agree:

Hypothesis 1a. When recruiters receive accurate and non-conflicting recommendations from both the human expert and the ADSS, the accuracy of their decision is higher.

Hypothesis 1b. When recruiters receive inaccurate and non-conflicting recommendations from both the human expert and the ADSS, the accuracy of their decision is lower.

In situations of disagreement between advisors, research shows that the resulting cognitive conflict of the judge may reduce the accuracy of their decision (Snizek & Buckley, 1995). In the case of personnel selection, studies tend to converge in favor of aversion behavior (Oberst et al., 2020). We therefore posit the following:

Hypothesis 2a. When recruiters receive conflicting recommendations, the accuracy of their decision is lower.

Hypothesis 2b. When recruiters receive conflicting recommendations, the influence of the human expert is higher than the ADSS.

Hypotheses related to post-task trust:

Drawing from experimental studies of decision making in complex tasks involving human and ADSS experts, we consider that the human expert is generally considered to be more trustworthy than an ADSS (Langer et al., 2022; Oberst et al., 2020; Rieger et al., 2022). We also consider that highly credible human experts may trigger an authority heuristic (Wang et al., 2020), especially when the decision making process is complex. We therefore posit the following:

Hypothesis 3. When recruiters receive non-conflicting recommendations, they consider recommendations from the human expert to be more trustworthy than ADSS recommendations.

Hypothesis 4. When recruiters receive conflicting recommendations, they consider accurate or inaccurate recommendations from a human expert to be more trustworthy than ADSS recommendations

METHODS

Our research design involved the participants (recruiters) collecting information and making their own impression about fictitious application forms. They were then exposed to a dual recommendation from a human expert and an ADSS, following which they were asked to make a final decision. This situation matches the usual practice among HRM practitioners whereby advice is used to complete their own evaluation. According to Bailey & al. (2022), this feature should maximize the influence of the recommendations on the recruiter's final choice. We manipulated two factors: the accuracy of the advice given (accurate vs. inaccurate), and the level of agreement between advisors (conflict vs. no conflict). We created a control group in which the recruiters received no advice, meaning they could be considered to be fully independent judges.

Development of study material

Firstly, we developed a fictitious job description for an HR Manager position inspired from those posted on the leading French job platform. Secondly, we created two resume abstracts inspired from real resumes in terms of layout and information provided (name, home and email address, qualifications, work experience, competency statement, and foreign language proficiency). In order to differentiate between the two job applications, we manipulated the amount of work experience, the academic background, the foreign language, and managerial skills. In this way, one of the two resumes fully met the person specification for the job, whereas the other resume only partially met the specification. To ensure their realistic nature and suitability, four HR management experts were asked to review the study material.

Thirdly, we generated five different recommendation conditions for each resume, as follows: accurate human expert / accurate ADSS (1); inaccurate human expert / inaccurate ADSS (2); inaccurate human expert / accurate ADSS (3); accurate human expert / inaccurate ADSS (4); no recommendation (5).

For the recommendation provided by the human expert, participants were given excerpts from phone conversations with a recruitment agency manager (the human expert), and they were informed that the agency had already been used several times for previous recruitments. It was also explained that the recruitment agency manager had received the job description and the two resumes before making a recommendation. The recommendations by the ADSS were based on a predictive recruitment solution developed by a fictional startup (HR Predict). Participants were informed that this startup had developed an advanced algorithmic solution, integrating the latest developments in AI (machine learning and deep learning). Participants were provided recommendations from HR Predict including a ranking of each applicant based on their match to the job description, and on their soft skills and technical skills in line with the job offer.

Participants

To enhance the ecological validity of our results, we selected professionals involved in the recruitment process in their organizations who were registered with a company specializing in panel surveys. We obtained 746 usable responses over a two-week period (147 in a first collection and then 599). The sample was composed of 47.7% male and 52.2% female participants, with an average age of 40.8 years ($SD = 10.23$). Participants were employed in companies ranging in size from 10 to over 1,000 employees (with a median class of 50-199). 51.1% of the respondents had three or more years' experience in recruiting, and 34.1% had one to 3 years' experience.

Procedure

Participants were initially asked to answer demographic questions. They were then required to read a job description for an HR Manager position, followed by two resume abstracts in response to the job offer. They were asked to assess each of the two abstracts on their own. To avoid order effects, the applicants' resumes were counterbalanced. After having read the job offer and the two abstracts, participants simultaneously received dual recommendations (from a human expert and an ADSS): they were randomly exposed to five different recommendation conditions, in parallel to a control group (in which participants were not exposed to any recommendations). As described above, the two manipulated factors were accuracy of recommendations (accurate vs. inaccurate) and the level of conflict (no conflict vs. conflict), meaning a disagreement between the recommendations (each recommendation identifying a different resume as the best).

As part of this filler task, participants were also required to answer questions developed to check their level of attention during the task and their ability to identify the differences between the two resumes. We excluded participants who did not successfully answer these questions. Finally, participants reported their level of trust in recommendations provided by the ADSS and the human expert. They were asked to decide which applicant they considered to be the best and should be invited to interview.

Measures

Dependent variables

We used two types of dependent variables:

- A variable reflecting recruiters' behavior: the percentage of most suitable resumes ranked in first position, reflecting participants accuracy of preselection decision. As previously mentioned, we considered a recommendation for the least suitable

application to be inaccurate, and a recommendation for the most suitable candidate to be accurate.

- Attitudinal variables reflecting recruiters' perceptions concerning their decision: self-confidence in own decision; trust in ADSS recommendation; trust in human expert recommendation.

We measured participants' trust in their own decisions using items proposed by Dietvorts (2016), scored on a 9-point Likert scale (e.g., *"How confident are you in your own decision?"*). We measured trust in the human expert or ADSS recommendations (Post-Task trust) with four items proposed by Dietvorst (2016) scored on a 9-point Likert scale: *"How trustworthy do you consider the recommendation provided by [the expert/the ADSS]?"*; *"How likely do you think it is that [the expert/the algorithmic solution] may fail ?"*. These measures assessed post-Task trust in the expert or automation, given that we took them after task completion. We also added two items asking participants to compare their decisions to ADSS or expert recommendations with two items proposed by Dietvorst (2016) and scored on a 9-point Likert scale: *"How trustworthy do you consider the recommendation provided by [the expert/the ADSS]?"*. Finally, using two binary-coded (yes/no) questions, we asked the respondents to indicate which advice they primarily relied on in making their decision (*"Did you primarily rely on [the expert/the ADSS] recommendation to make your decision?"*).

Control variables

Hoff and Bashir (2015) showed that age can have an effect on trust in automation, as older adults tend to comply with and rely on automation more than younger adults. Related to gender differences, studies are still scarce, and their results remain inconsistent (Hoff and Bashir, 2015). Consequently, we decided to collect sociodemographic information (*age* and *gender*) to be used as control variables.

Research conducted on trust in algorithmic systems concludes that personal characteristics - especially self-confidence and propensity to trust automation - has an influence on actual trust in recommendations provided by such systems (Jessup et al., 2019; J. D. Lee & See, 2004). To measure participants' *self-confidence in their ability to recruit*, we adapted the Williams and Deci (1996) perceived competence scale (4 items); a sample item being: "I feel confident in my ability to conduct a successful recruitment" (Cronbach's $\alpha = 0.91$). We assessed *propensity to trust automated systems* using the 9-item scale of trust between people and automation adapted from Jian et al. (2000); two sample 7-point Likert items being: "ADSS are deceptive" (reverse-coded) and "ADSS are reliable" (Cronbach's $\alpha = 0.84$).

Research has also shown that the *recruiters' expertise in recruitment* may also influence their attitudes and behavior. We therefore included the two following questions: "How many years have you been involved in these selection/recruitment procedures?" and "On average, during each of these selection procedures, how many applications do you screen?" We computed a composite score representing expertise in recruitment as $\log(\text{experience} \times \text{number of applications screened})$. We checked the participants' former use of ADSS by asking them if they had already *used an ADSS* in their professional practice, and recoded as a binary item (yes/no).

Results

Descriptive statistics and manipulation checks

Checking the control condition shows that respondents distinguished one resume from the other: the most suitable (the best resume) was ranked in first place in 64% of cases compared to 35% for the second resume. The overall accuracy of the control group could therefore be evaluated at 64%.

The experimental conditions showed differentiated impacts on respondents' choices and trust (recruiter's trust in external advisors and recruiter's trust in their own opinion, see Table 1 & Table 2). Chi-squared test of independence between the displayed conditions is highly significant (Chi-squared (4, N=746) = 40.71, $p < .001$, Cramer's V = 0.23). ANOVA conducted on the five dependent variables concerning trust were significant, except for the variable "compare self-expert". It can be concluded that the experimental material was of sufficient quality to address the hypotheses we developed.

 Insert Table 1 about here

 Insert Table 2 about here

Testing of hypotheses

Regarding the hypotheses on recruiters' preselection decisions, H1a. stated that when recruiters received accurate and non-conflictual recommendations from both human experts and ADSS, the accuracy of their decision would be higher. Results showed that accurate recommendations provided by both human experts and ADSS significantly increased the proportion of best resumes ranked in first position from 64.63% in the control group to 79.87% in the experimental group, in support of H1a: Chi-squared (1, N=296) = 8.58, $p = .004$, Cramer's V = 0.17, Odds Ratio = 2.17 [1.29, 3.67]).

H1b. stated that when recruiters received inaccurate and non-conflictual recommendations from both the human expert and the ADSS, the accuracy of their decision would be lower (compared to the control conditions). Results showed that inaccurate and non-conflicting recommendations provided by both human experts and ADSS significantly decreased the proportion of best resumes ranked in first position from 64.63% in the control group to 48.32 % in the

experimental group: Chi-squared (1, N=296) = 8, $p = .004$, Cramer's $V = 0.16$, Odds Ratio = 0.51 [0.32, 0.82]). Inaccurate recommendations from both a human expert and an ADSS divided by 1.96 the odds of the most suitable resume being ranked in first position. Therefore H1b. is supported. The extent of the effect is almost the same as in congruent and accurate recommendations, but in the opposite direction. It appeared that recommendations had a very similar influence when both human experts and ADSS gave the same advice, regardless of the quality of this.

To further investigate the effect of dual recommendations provided to recruiters, we proposed hypotheses H2a and H2b. H2a posited that when recruiters received conflicting recommendations, the accuracy of their decision would be lower, in line with results obtained by Snizek and Buckley (1995). Results showed a statistically significant decrease in the average accuracy of the recruiters' decision confronted with conflicting recommendations as compared with the control group: the rate of correctly ranked resumes decreased by 10 points from 64.6 % to 54.6 % (Chi-squared (1, N= 448) = 3.90, $p = 0.04$, Cramer's $V = 0.09$, Odds Ratio = 0.66 [0.44; 1]). Conflicting recommendations divided by 1.5 the odds of the best resume being ranked in first position. Our data therefore support H2a. in line with Snizek & Buckley's (1995) experimental results. Hypothesis H2b. posited that when recruiters received conflicting recommendations, the average influence of the human expert would be higher than the ADSS on the recruiters' decision. This should result in superior accuracy for the condition "accurate expert/inaccurate ADSS" compared with the condition "inaccurate expert/accurate ADSS". Results showed the opposite: decision accuracy is higher for the condition "inaccurate expert/accurate ADSS" (59.7%) compared with the condition "accurate expert/inaccurate ADSS" (50%). This result is statistically non-significant: Chi-squared (1, N=301) = 2.88, $p = .09$, Cramer's $V = 0.1$, Odds ratio = 1.48 [0.94;2.34]. In this situation of cognitive conflict for

recruiters, they slightly favored the advice given by the ADSS, leading us to conclude that H2.b is not supported.

Concerning the hypotheses related to post-task trust, H3 posited that when recruiters received non-conflicting recommendations from both human expert and ADSS, they would consider recommendations from a human expert to be more trustworthy than ADSS recommendations. Paired samples T-tests (Table 3) revealed that in these cases recruiters had more trust in the human expert than the ADSS. H3 is therefore supported.

 Insert Table 3 about here

Concerning the hypotheses related to disagreement between advisors leading to cognitive conflict, we proposed H4. positing that in situations of disagreement between advisors (conflict between the recommendations made by human experts and the ADSS), human recommendation would be considered to be more trustworthy than that of an ADSS, all things being equal. As shown in table 3, paired samples T-tests revealed that recruiters considered human experts to be more trustworthy, even when the expert is inaccurate. H4 is therefore supported.

Cases of cognitive conflict are particularly useful for comparing the relative impact of each item of advice. In such situations, the expert is considered to be more trustworthy, but it should be remembered that ADSS was more influential on the recruiter's behavior. This “paradox of trust” is also illustrated by the perceived probability of failure reported in Table 3: in non-conflicting situations, no differences were found in perceived probability that either the ADSS or the human expert could give inaccurate advice. On the opposite, in conflicting situations recruiters considered that the probability of failure is always higher for ADSS, even if the human expert gave inaccurate advice; this perception follows the overall lack of declared trust in ADSS.

Additional results: predicting that advice will be followed

Since hypothesis H2b was not supported, we decided to explore the factors influencing the decision to follow the recommendations provided by either ADSS or the human expert in conflicting situations, including the control variables. Additional analyses were undertaken on the subsample that addressed these conditions ($N = 301$), in order to isolate the predictors that led to one of the two opposing recommendations being followed.

Cross-tabulation results obtained from the repeated measured variable “decision to follow expert and/or ADSS recommendations” showed that recruiters declared they were more likely to follow the advice provided by the human expert in conflicting situations. There were more discordant "expert followed/ADSS not followed" pairs (97) than discordant "ADSS followed/expert not followed" pairs (45), with the difference being statistically significant (McNemar Chi-squared (1, $N=301$) = 18.31, $p < .001$, Odds ratio = 2.15 [1.49;3.14]). This result confirmed the findings of the hypothesis tests, with participants claiming that the human expert was more trustworthy, and reporting that they more often aligned themselves with human advice (the odds being 2.5 times higher).

In the next stage, we ran two logistic regression models incorporating socio-demographic predictors, reported trust and comparisons, perceived competence, and personality traits (perceived self-competence, propensity to trust automation).

 Insert Table 4 about here

Results (table 4) showed some differences in predictors for the two models. The decision to follow expert recommendation (model 1) was mainly driven by trust in experts, and negatively associated with self confidence in one's own evaluation, which may indicate a possible

authority heuristic (Wang et al., 2020). We noticed a paradox, namely that when recruiters considered themselves to be better than experts, they more often declare following their advice.

Concerning the decision to comply with ADSS recommendations (model 2), our results showed that trust in ADSS and propensity to trust automation had a significant and positive influence in line with previous studies (Jessup et al., 2019; J. D. Lee & See, 2004). We noticed that when recruiters do not feel very confident in their ability to recruit, and when they tend to value human experts more than themselves, they are more prone to follow ADSS recommendations. Our results highlighted the same paradox as illustrated just above, namely that when recruiters found themselves to be better than ADSS, they tended to comply more with automated recommendations. Recruiter expertise increased the tendency to follow ADSS recommendations, but not expert recommendations. We also noticed a discrepancy between reported and measured level of expertise in model 2: auto-reported self confidence in recruitment negatively influences the decision to follow ADSS, while recruiter ‘expertise’ positively influenced the decision to follow ADSS. Age played a marginal role: older recruiters were slightly more reluctant to follow both expert and ADSS recommendations.

Discussion

Our study provides two main findings: 1) Non-human and human experts are highly influential in recruitment decision-making processes but not equally trusted, and 2) the advice of human experts is trusted but not followed, especially in situations of conflictual advice.

Does congruence lead to conformity?

The first noticeable finding from our study is the influence of congruent advice on the recruiter’s decision. On the one hand, accurate recommendations provided by both a human expert and an ADSS resulted in an improvement in the accuracy of recruiters’ decisions. In line with Snizek and Buckley (1995), we saw that accurate advice was, as expected, beneficial. On the other

hand, when both ADSS and human experts provided inaccurate recommendations, the accuracy of recruiter's decisions was reduced to a similar extent. Recruiters followed accurate and inaccurate advice to the same extent (about 10 percentage points above or below the control group). The recruiter's own expertise did not appear to interfere with the accuracy of the recommendation. We suggest two explanations for this:

The first is methodological: for ecological purposes, differences between resumes were small in our experimental material. The situation was somewhat ambiguous, so that even experienced recruiters may have found relevant or acceptable an inaccurate but congruent recommendation (where human experts and ADSS both suggested to choose the second resume).

The second explanation is related to the concept of conformity in social psychology. Although the situation was very different, we could invoke a conformity effect, as revealed by a long series of experiments conducted following Solomon Ash's (1956) seminal work. Ash demonstrated that the unanimity in a group of respondents may exert an influence on a lonely subject's own choice, even when the suggested decision is obviously inaccurate. Unanimity of advisors might have had a major influence on the recruiter's choice in our sample. Studies attempting to replicate Ash's experiments on conformity in situations including non-human advisors led to inconclusive results: some studies showed that, unlike human peers, robot peers cannot create conformity based on social influence (Brandstetter et al., 2014; Shiomi & Hagita, 2016), while other studies demonstrated that conformity may occur when participants are allowed to assess the accuracy of robot advice before making their own decision (Salomons et al., 2018). A recent study (Masjutin et al., 2022) replicating the experiment by Ash with a "hybrid majority" composed of humans and robots showed that participants tended to conform most often with hybrid majorities (two humans and two robots) compared with robot majorities (four robots and one human). This result is informative, but we must keep in mind that the size of our hybrid group (two advisors) is smaller than those used in earlier cited studies conducted

within Ash's experimental framework and does not reach the threshold for exerting social influence (three to four confederates in Ash's original experiment).

The trust paradox: why rely on machine heuristics and say the opposite?

The second main finding from our study is the discrepancy between reported trust and actual recruiter behaviors when they were exposed to inaccurate ADSS advice. In case of disagreement between the recommendations by human expert and the ADSS, we observed that the quality of recruiters' decisions decreased. This result is in line with the work of Snizek and Buckley (1995) who found, in contradiction to their initial hypothesis, that cognitive conflict led the subjects to reduce their cognitive effort, considering that the disagreement between advisors may be a cue indicating that the problem to be solved is difficult. Instead of increasing their accuracy by processing more information and deepening their analysis, they may have used heuristics. In the present case, we suggest a machine heuristic, as described by Sundar & Kim (2019).

Compared with the recommendation made by the human expert, that provided by the ADSS was considered to not be as equally trustworthy, and more prone to failures. However, in the meantime it was influential on the decision making of recruiters in conflicting conditions. The most illustrative case is the "expert accurate / ADSS inaccurate" condition, in which the accuracy of the decision was reduced by 14 points (from 64.6% to 50%) compared to the control group. The difference is statistically significant ($\chi^2(1, N=299) = 6.53, p = 0.01$, Odds ratio = 1.83). Inaccurate recommendations provided by the human expert also had an impact: the accuracy of the decision falls from 64% in the control group to 59%, with a smaller and non-significant effect ($\chi^2(1, N=296) = 0.75, p = 0.39$).

The question is therefore, if the recruiters follow the ADSS in making their decision, why do they not say so? In other words, the choice to rely on ADSS advice is not overtly assumed, as

respondents state after making their decision that they trust the human expert more than the ADSS. This attitude-behavior inconsistency reflects an important stream of research in social psychology (Fazio & Zanna, 1981; Kraus, 1995; McBroom & Reed, 1992). Several variables influencing this inconsistency have been proposed, from individual predictors, such as self-monitoring (see DeBono & Omoto, 1993) to situational variables (presence of others, lack of knowledge). Another explanation for the discrepancy between what recruiters claim about their behavior and their actual decision, is that the question of AI in personnel selection remains a sensitive issue (in particular for fear of being replaced by superior machines). When the topic is sensitive it has been demonstrated that there is no straightforward relationship between people's actions and attitudes especially because of social desirability pressures (Pager & Quillian, 2005; Wulff & Villadsen, 2020). Some research also highlighted a tendency for cognitive laziness when people received help from AI because of a reluctance to engage in a cognitively demanding mental process. This can explain the observed compliance with the ADSS recommendation (Skitka et al., 2000).

Overall, our empirical results raise some managerial issues given that the consequences of activating a machine heuristic are potentially detrimental. Algorithms are not perfectly reliable in the recruitment domain because of technical errors and the reproduction of discriminations. Köchling et al., (2021) have shown that predictive recruitment models reproduce the biases existing in the training data, and even tend to amplify them, and disadvantage the categories that are under-represented in the data (particularly on the basis of ethnicity). If recruiters blindly follow the ADSS, they may reproduce discriminations. As a consequence, keeping humans in the loop as generally suggested in studies based on a cooperative framework (Alon-Barkat & Busuioc, 2021) is not enough if there is a machine heuristic and if recommendations based on AI are not perfectly reliable. Further efforts based on a systematic audit of the reliability and effectiveness of ADSS are necessary to address ADSS biases that can be substantiated by

human recruiters subject to machine heuristics (Hunkenschroer & Luetge, 2022; Kazim et al., 2021).

Limitations and further research opportunities

Certain limitations remain, and these should be addressed in future research. Firstly, isolating the unique influence of recommendations in recruiter choices was difficult, given that the task assigned in this study was complex and subjective. Recruiter choices may have reflected a certain subjectivity (as shown in the control group in which the least suitable resume was selected by 35% of respondents). Secondly, our concern for ecological validity led us to create credible application forms, whose characteristics were consistent with the job offer, and which were not strongly differentiated: this led to small effect sizes in between-group comparisons.

Decision making in augmented recruitment is a cognitive task: situational variables identified in the literature on trust in automation such as workload or time constraints could be considered to further inform the antecedents of recruiter choice to rely or not on ADSS (J. D. Lee & See, 2004; Solberg et al., 2022). Our results finally suggest that future research is needed to address the issue of potential social desirability bias in recruiters' responses, in line with the increasing use of AI in complex and subjective stages of the selection process.

Conclusion

In a context of the generalization of augmented recruitment using human-AI collaboration in decision making tasks and subsequently in the development of hybrid decision making processes, we aimed to explore the mechanisms underlying recruiters' trust and behavior towards both human expert and ADSS recommendations in supporting their preselection decisions. Our experiment explored the source, the accuracy and the congruence of the recommendations provided. Our results suggest that recruiters are strongly influenced by the advice provided by a dual recommendation. They are more influenced by advice from ADSS

than a human expert, even if they declare the human expert to be more trustworthy than ADSS. Such findings raise questions regarding the cognitive process that leads recruiters to rely on algorithmic advice in a complex and subjective task such as the screening of resumes. Our research points to the need for caution when a recruiter is exposed to algorithmic recommendations, even if a peer also provides advice at the same time.

References

- Ash, S.E. (1956). Studies of Independence and Conformity, "A Minority of One Against a Unanimous Majority", *Psychological Monograph: General and Applied*, 70, Whole No. 416..
- Alon-Barkat, S., & Busuioc, M. (2021). Decision-makers Processing of AI Algorithmic Advice: Automation Bias versus Selective Adherence. *arXiv preprint* arXiv:2103.02381.
- Bailey, P. E., Leon, T., Ebner, N. C., Moustafa, A. A., & Weidemann, G. (2022). A meta-analysis of the weight of advice in decision-making. *Current Psychology*. <https://doi.org/10.1007/s12144-022-03573-2>
- Bogert, E., Schechter, A., & Watson, R. T. (2021). Humans rely more on algorithms than social influence as a task becomes more difficult. *Scientific Reports*, 11(1), Art. 1. <https://doi.org/10.1038/s41598-021-87480-9>
- Brandstetter, J., Rácz, P., Beckner, C., Sandoval, E. B., Hay, J., & Bartneck, C. (2014). A peer pressure experiment: Recreation of the Asch conformity experiment with robots. 2014 *IEEE/RSJ International Conference on Intelligent Robots and Systems*, 1335-1340. <https://doi.org/10.1109/IROS.2014.6942730>
- Brynjolfsson, E., & Mitchell, T. (2017). What can machine learning do? Workforce implications. *Science*, 358(6370), 1530-1534. <https://doi.org/10.1126/science.aap8062>
- Chamorro-Premuzic, T., Winsborough, D., Sherman, R. A., & Hogan, R. (2016). New Talent Signals: Shiny New Objects or a Brave New World? *Industrial and Organizational Psychology*, 9(3), 621-640. <https://doi.org/10.1017/iop.2016.6>
- De Cremer, D. (2020). *Leadership by algorithm: Who leads and who follows in the AI era?* Harriman House Limited.
- DeBono, K. G., & Omoto, A. M. (1993). Individual Differences in Predicting Behavioral Intentions from Attitude and Subjective Norm. *Journal of Social Psychology*, 133(6), 825-831. <https://doi.org/10.1080/00224545.1993.9713944>

- Dietvorst, B. J., Simmons, J. P., & Massey, C. (2015). Algorithm aversion: People erroneously avoid algorithms after seeing them err. *Journal of Experimental Psychology: General*, 144(1), 114-126. doi: 10.1037/xge0000033
- Dzindolet, M. T., Pierce, L. G., Beck, H. P., & Dawe, L. A. (2002). The perceived utility of human and automated aids in a visual detection task. *Human Factors*, 44, 79-94. <https://doi.org/10.1518/0018720024494856>
- Fazio, R. H., & Zanna, M. P. (1981). Direct experience and attitude-behavior consistency. In *Advances in experimental social psychology*, 14, 61-202. [https://doi.org/10.1016/S0065-2601\(08\)60372-X](https://doi.org/10.1016/S0065-2601(08)60372-X)
- Gillath, O., Ai, T., Branicky, M. S., Keshmiri, S., Davison, R. B., & Spaulding, R. (2021). Attachment and trust in artificial intelligence. *Computers in Human Behavior*, 115, 106607. <https://doi.org/10.1016/j.chb.2020.106607>
- Glikson, E., & Woolley, A. W. (2020). Human Trust in Artificial Intelligence: Review of Empirical Research. *Academy of Management Annals*, 14(2), 627-660. <https://doi.org/10.5465/annals.2018.0057>
- Gonzalez, M., F., Capman, J., F., Oswald, F. L., They, E., R., & Tomczak, D., L. (2019). “Where’s the I-O?” Artificial Intelligence and Machine Learning in Talent Management Systems. *Personnel Assessment and Decisions*, 5(3), 33-44. <https://doi.org/10.25035/pad.2019.03.005>
- Grønsund, T., & Aanestad, M. (2020). Augmenting the algorithm: Emerging human-in-the-loop work configurations. *The Journal of Strategic Information Systems*, 29(2), 101614. <https://doi.org/10.1016/j.jsis.2020.101614>
- Groom, V., & Nass, C. (2007). Can robots be teammates?: Benchmarks in human–robot teams. *Interaction Studies*, 8(3), 483-500. <https://doi.org/10.1075/is.8.3.10gro>
- Grove, W. M., Zald, D. H., Lebow, B. S., Snitz, B. E., & Nelson, C. (2000). Clinical versus mechanical prediction: A meta-analysis. *Psychological assessment*, 12(1), 19-30. <http://doi.apa.org/getdoi.cfm?doi=10.1037/1040-3590.12.1.19>
- Haesevoets, T., De Cremer, D., Dierckx, K., & Van Hiel, A. (2021). Human-machine collaboration in managerial decision making. *Computers in Human Behavior*, 119, 106730. <https://doi.org/10.1016/j.chb.2021.106730>
- Harvey, N., & Fischer, I. (1997). Taking Advice: Accepting Help, Improving Judgment, and Sharing Responsibility. *Organizational Behavior and Human Decision Processes*, 70(2), 117-133. <https://doi.org/10.1006/obhd.1997.2697>
- Hoff, K. A., & Bashir, M. (2015). Trust in Automation: Integrating Empirical Evidence on Factors That Influence Trust. *Human Factors*, 57(3), 407-434. <https://doi.org/10.1177/0018720814547570>

- Hou, Y. T.-Y., & Jung, M. F. (2021). Who is the expert? Reconciling algorithm aversion and algorithm appreciation in AI-supported decision making. *Proceedings of the ACM on Human-Computer Interaction*, 5(CSCW2), 1-25. <https://doi.org/10.1145/3479864>
- Hunkenschroer, A. L., & Luetge, C. (2022). Ethics of AI-Enabled Recruiting and Selection: A Review and Research Agenda. *Journal of Business Ethics*, 178(4), 977-1007. <https://doi.org/10.1007/s10551-022-05049-6>
- Jessup, S.A., Schneider, T.R., Alarcon, G.M., Ryan, T.J., Capiola, A. (2019). The Measurement of the Propensity to Trust Automation. In: Chen, J., Fragomeni, G. (eds) *Virtual, Augmented and Mixed Reality. Applications and Case Studies*. HCII 2019. Lecture Notes in Computer Science, vol 11575. Springer, Cham. https://doi.org/10.1007/978-3-030-21565-1_32
- Kazim, E., Denny, D. M. T., & Koshiyama, A. (2021). AI auditing and impact assessment: According to the UK information commissioner's office. *AI and Ethics*, 1(3), 301-310. <https://doi.org/10.1007/s43681-021-00039-2>
- Köchling, A., Riazzy, S., Wehner, M. C., & Simbeck, K. (2021). Highly Accurate, But Still Discriminatory. *Business & Information Systems Engineering*, 63(1), 39-54. <https://doi.org/10.1007/s12599-020-00673-w>
- König, C. J., & Langer, M. (2022). Machine learning in personnel selection. In S. Strohmeier (dir) *Handbook of research on Artificial Intelligence in Human Resource Management*. Edward Elgar Publishing, 149-167
- Kraus, S. J. (1995). Attitudes and the Prediction of Behavior: A Meta-Analysis of the Empirical Literature. *Personality and Social Psychology Bulletin*, 21(1), 58-75. <https://doi.org/10.1177/0146167295211007>
- Kuncel, N. R., Klieger, D. M., Connelly, B. S., & Ones, D. S. (2013). Mechanical versus clinical data combination in selection and admissions decisions: A meta-analysis. *Journal of Applied Psychology*, 98(6), 1060-1072. <https://psycnet.apa.org/doi/10.1037/a0034156>
- Langer, M., König, C., Back, C., & Hemsing, V. (2022). Trust in Artificial Intelligence: Comparing trust processes between human and automated trustees in light of unfair bias. *Journal of Business and Psychology*. <https://doi.org/10.31234/osf.io/r9y3t>
- Langer, M., König, C. J., & Busch, V. (2021). Changing the means of managerial work: Effects of automated decision support systems on personnel selection tasks. *Journal of Business and Psychology*, 36(5), 751-769. <https://doi.org/10.1007/s10869-020-09711-6>
- Langer, M., König, C. J., Sanchez, D. R.-P., & Samadi, S. (2019). Highly automated interviews: Applicant reactions and the organizational context. *Journal of Managerial Psychology*, 35(4), 301-314

- Langer, M., & Landers, R. N. (2021). The future of artificial intelligence at work: A review on effects of decision automation and augmentation on workers targeted by algorithms and third-party observers. *Computers in Human Behavior*, 123, 106878. <https://doi.org/10.1016/j.chb.2021.106878>
- Lee, J. D., & See, K. A. (2004). Trust in Automation: Designing for Appropriate Reliance. *Human Factors*, 46(1), 50-80. https://doi.org/10.1518/hfes.46.1.50_30392
- Lee, M. K. (2018). Understanding perception of algorithmic decisions: Fairness, trust, and emotion in response to algorithmic management. *Big Data & Society*, 5(1), 2053951718756684. <https://doi.org/10.1177/2053951718756684>
- Liel, Y., & Zalmanson, L. (2020). What If an AI Told You That $2 + 2$ Is 5? Conformity to Algorithmic Recommendations. *International Conference on Information Systems, ICIS 2020 - Making Digital Inclusive: Blending the Local and the Global*. ISBN: 9781733632553
- Logg, J. M., Minson, J. A., & Moore, D. A. (2019). Algorithm appreciation: People prefer algorithmic to human judgment. *Organizational Behavior and Human Decision Processes*, 151, 90-103. <https://doi.org/10.1016/j.obhdp.2018.12.005>
- Madhavan, P., & Wiegmann, D. A. (2007). Similarities and differences between human-human and human-automation trust: An integrative review. *Theoretical Issues in Ergonomics Science*, 8(4), 277-301. <https://doi.org/10.1080/14639220500337708>
- Masjutin, L., Laing, J. K., & Maier, G. W. (2022). Why do We Follow Robots? An Experimental Investigation of Conformity with Robot, Human, and Hybrid Majorities. *17th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, 139-146. <https://doi.org/10.1109/HRI53351.2022.9889675>
- Mayer, R. C., Davis, J. H., & Schoorman, F. D. (1995). An Integrative Model of Organizational Trust. *Academy of Management Review*, 20(3), 709-734. <https://doi.org/10.5465/amr.1995.9508080335>
- McBroom, W. H., & Reed, F. W. (1992). Toward a Reconceptualization of Attitude-Behavior Consistency. *Social Psychology Quarterly*, 55(2), 205-216. <https://doi.org/10.2307/2786946>
- Mesbah, N., Tauchert, C., & Buxmann, P. (2021). Whose advice counts more—man or machine? An experimental investigation of ai-based advice utilization. Online Conference, *54th Hawaii International Conference on System Sciences*, 05.08.01.2021,
- Nawaz, N. (2020). Artificial Intelligence Applications for Face Recognition in Recruitment Process. *Journal of Management Information & Decision Sciences*, 23, 499-509.
- Oberst, U., De Quintana, M., Del Cerro, S., & Chamarro, A. (2020). Recruiters prefer expert recommendations over digital hiring algorithm: A choice-based conjoint study in a

- pre-employment screening scenario. *Management Research Review*, 44(4), 625-641.
doi: 10.1108/MRR-06-2020-0356
- Pager, D., & Quillian, L. (2005). Walking the Talk? What Employers Say versus What They Do. *American Sociological Review*, 70(3), 355-380.
<https://doi.org/10.1177/000312240507000301>
- Pescetelli, N., & Yeung, N. (2021). The role of decision confidence in advice-taking and trust formation. *Journal of Experimental Psychology: General*, 150, 507-526.
<https://doi.org/10.1037/xge0000960>
- Prikshat, V., Malik, A., & Budhwar, P. (2023). AI-augmented HRM: Antecedents, assimilation and multilevel consequences. *Human Resource Management Review*, 33(1), 100860. <https://doi.org/10.1016/j.hrmr.2021.100860>
- Raisch, S., & Krakowski, S. (2021). Artificial intelligence and management: The automation–augmentation paradox. *Academy of Management Review*, 46(1), 192-210.
<https://doi.org/10.5465/amr.2018.0072>
- Rieger, T., Roesler, E., & Manzey, D. (2022). Challenging presumed technological superiority when working with (artificial) colleagues. *Scientific Reports*, 12(1), Art. 1.
<https://doi.org/10.1038/s41598-022-07808-x>
- Salomons, N., van der Linden, M., Strohkorb Sebo, S., & Scassellati, B. (2018). Humans Conform to Robots: Disambiguating Trust, Truth, and Conformity. *Proceedings of the 2018 ACM/IEEE International Conference on Human-Robot Interaction*, 187-195.
<https://doi.org/10.1145/3171221.3171282>
- Schemmer, M., Hemmer, P., Kühl, N., Benz, C., & Satzger, G. (2022). Should I Follow AI-based Advice? Measuring Appropriate Reliance in Human-AI Decision-Making. *arXiv preprint arXiv:2204.06916*.doi: 10.48550/arXiv.2204.06916
- Shiomi, M., & Hagita, N. (2016). Do Synchronized Multiple Robots Exert Peer Pressure? *Proceedings of the Fourth International Conference on Human Agent Interaction*, 27-33. <https://doi.org/10.1145/2974804.2974808>
- Shrestha, Y. R., Ben-Menahem, S. M., & Von Krogh, G. (2019). Organizational decision-making structures in the age of artificial intelligence. *California Management Review*, 61(4), 66-83. <https://doi.org/10.1177/0008125619862257>.
- Skitka, L. J., Mosier, K., & Burdick, M. D. (2000). Accountability and automation bias. *International Journal of Human-Computer Studies*, 52(4), 701-717.
<https://doi.org/10.1006/ijhc.1999.0349>
- Snizek, J. A., & Buckley, T. (1995). Cueing and Cognitive Conflict in Judge-Advisor Decision Making. *Organizational Behavior and Human Decision Processes*, 62(2), 159-174. <https://doi.org/10.1006/obhd.1995.1040>

- Snizek, J. A., Paese, P. W., & Switzer, F. S. (1990). The effect of choosing on confidence in choice. *Organizational Behavior and Human Decision Processes*, 46(2), 264-282. [https://doi.org/10.1016/0749-5978\(90\)90032-5](https://doi.org/10.1016/0749-5978(90)90032-5)
- Snizek, J. A., & Van Swol, L. M. (2001). Trust, Confidence, and Expertise in a Judge-Advisor System. *Organizational Behavior and Human Decision Processes*, 84(2), 288-307. <https://doi.org/10.1006/obhd.2000.2926>
- Solberg, E., Kaarstad, M., Eitrheim, M. H. R., Bisio, R., Reegård, K., & Bloch, M. (2022). A Conceptual Model of Trust, Perceived Risk, and Reliance on AI Decision Aids. *Group & Organization Management*, 47(2), 187-222. <https://doi.org/10.1177/10596011221081238>
- Sowa, K., Przegalinska, A., & Ciechanowski, L. (2021). Cobots in knowledge work: Human – AI collaboration in managerial professions. *Journal of Business Research*, 125, 135-142. <https://doi.org/10.1016/j.jbusres.2020.11.038>
- Sundar, S. S., & Kim, J. (2019). Machine Heuristic: When We Trust Computers More than Humans with Our Personal Information. *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, 1-9. <https://doi.org/10.1145/3290605.3300768>
- Wang, J., Molina, M. D., & Sundar, S. S. (2020). When expert recommendation contradicts peer opinion : Relative social influence of valence, group identity and artificial intelligence. *Computers in Human Behavior*, 107, 106278. <https://doi.org/10.1016/j.chb.2020.106278>
- Wulff, J. N., & Villadsen, A. R. (2020). Are Survey Experiments as Valid as Field Experiments in Management Research? An Empirical Comparison Using the Case of Ethnic Employment Discrimination. *European Management Review*, 17(1), 347-356. <https://doi.org/10.1111/emre.12342>
- Yaniv, I. (2004). Receiving other people's advice: Influence and benefit. *Organizational Behavior and Human Decision Processes*, 93(1), 1-13. <https://doi.org/10.1016/j.obhdp.2003.08.002>
- You, S., Yang, C. L., & Li, X. (2022). Algorithmic versus Human Advice: Does Presenting Prediction Performance Matter for Algorithm Appreciation? *Journal of Management Information Systems*, 39(2), 336-365. <https://doi.org/10.1080/07421222.2022.2063553>

Tables

Table 1: Decision made under experimental conditions

Conditions		CV ranked 1	
		Least suitable CV	Most suitable CV
Control (no recommendation)	N	52	95
	%	35.37 %	64.63 %
Congruent and accurate recommendations	N	30	119
	%	20.13 %	79.87 %
Congruent and inaccurate recommendations	N	77	72
	%	51.68 %	48.32 %
ADSS accurate / Expert inaccurate	N	60	89
	%	40.27 %	59.73 %
ADSS inaccurate / Expert accurate	N	76	76
	%	50.00 %	50.00 %

Chi-squared (4, N=746) = 40.71, $p < .001$, Cramer's $V = 0.23$

Table 2. Descriptive statistics on recruiter's post-task trust and confidence

Conditions	Mean (sd) of recruiter's trust in external advisors		Mean (sd) of recruiter's trust in their own opinion		
	Trust ADSS	Trust expert	Self confidence	Compare Self vs ADSS	Compare Self vs expert
ADSS & Expert accurate	6.85 (1.25)	7.21 (1.12)	7.36 (1.05)	5.87 (1.33)	6.05 (1.34)
ADSS & Expert inaccurate	6.50 (1.27)	6.93 (1.07)	7.24 (0.94)	5.90 (1.46)	6.17 (1.24)
ADSS accurate Expert inaccurate	6.58 (1.30)	7.30 (0.96)	7.44 (1.02)	5.96 (1.51)	6.24 (1.38)
ADSS inaccurate Expert accurate	6.44 (1.34)	7.14 (1.02)	7.14 (1.03)	5.45 (1.44)	5.99 (1.28)
Overall Mean	6.59 (0.18)	7.14 (0.15)	7.29 (0.13)	5.79 (0.23)	6.11 (0.11)
ANOVA	F(3,595) = 3.00*	F(3,595) = 3.26*	F(3,595) = 2.42*	F(3,595) = 4.02**	F(3,595) = 1.06 ns

Table 3. Paired sample T-tests on perceived trust and perceived probability of failure

Conditions	Comparisons	Mean	CI	CI	<i>p</i>	Effect size d
		diff.	lower	upper		
ADSS & Expert accurate	trust_ADSS - trust_expert	-.356	-.573	-.139	.001	-0.27
	prob_ADSS_fail - prob_expert_fail	.168	-.050	.385	.129	0.12
ADSS & Expert inaccurate	trust_ADSS - trust_expert	-.436	-.657	-.216	.000	-0.32
	prob_ADSS_fail - prob_expert_fail	.161	-.110	.433	.243	0.10
ADSS accurate Expert inaccurate	trust_ADSS - trust_expert	-.718	-.926	-.510	.000	-0.56
	prob_ADSS_fail - prob_expert_fail	.356	.078	.633	.012	0.21
ADSS inaccurate Expert accurate	trust_ADSS - trust_expert	-.697	-.920	-.474	.000	-0.50
	prob_ADSS_fail - prob_expert_fail	.263	.032	.494	.026	0.18

Table 4. Logistic regression models on following of recommendations

	Model 1				Model 2			
	DV = expert advice followed				DV = ADSS advice followed			
	Estimate	SE	OR	<i>p</i>	Estimate	SE	OR	<i>p</i>
Trust own evaluation	-0.25	0.11	0.78	0.03	-0.08	0.12	0.92	0.49
Trust ADSS recommendation	-0.2	0.11	0.82	0.06	0.39	0.12	1.47	< .001
Trust expert recommendation	0.24	0.11	1.27	0.03	-0.05	0.12	0.96	0.71
Compare self vs expert	0.31	0.1	1.37	< .001	-0.27	0.1	0.76	0.01
Compare self vs ADSS	-0.12	0.1	0.89	0.24	0.28	0.11	1.32	0.01
Propensity to trust automation	-0.02	0.11	0.98	0.82	0.74	0.13	2.09	< .001
Self-competence	0.11	0.11	1.12	0.31	-0.24	0.12	0.79	0.04
Age	-0.03	0.01	0.97	< .01	-0.05	0.01	0.95	< .001
Algo use	-0.2	0.11	0.82	0.06	-0.14	0.11	0.87	0.19
Gender (female)	-0.36	0.19	0.7	0.06	-0.11	0.2	0.9	0.59
Recruiter's expertise	-0.03	0.13	0.97	0.79	0.35	0.13	1.43	0.01
R ² (Nagelkerke)	0.11				0.25			

Note. When a recommendation provided by either ADSS or the expert was followed, it was coded as class 1.