ORIGINAL PAPER

Changing the means of managerial work: effects of automated decision support systems on personnel selection tasks

Markus Langer[1] · Cornelius J. König[1] · Vivien Busch[1]

Abstract
To enhance the quality and efficiency of information processing and decision-making, automation based on artificial intelligence and machine learning has increasingly been used to support managerial tasks and duties. In contrast to classical applications of automation (e.g., within production or aviation), little is known about how the implementation of automation for management changes managerial work. In a work design frame, this study investigates how different versions of automated decision support systems for personnel selection as a specific management task affect decision task performance, time to reach a decision, reactions to the task (e.g., enjoyment), and self-efficacy in personnel selection. In a laboratory experiment, participants ($N = 122$) were randomly assigned to three groups and performed five rounds of a personnel selection task. The first group received a ranking of the applicants by an automated support system *before* participants processed applicant information (support-before-processing group), the second group received a ranking *after* they processed applicant information (support-after-processing group), and the third group received no ranking (no-support group). Results showed that satisfaction with the decision was higher for the support-after-processing group. Furthermore, participants in this group showed a steeper increase in self-efficacy in personnel selection compared to the other groups. This study combines human factors, management, and industrial/organizational psychology literature and goes beyond discussions concerning effectiveness and efficiency in the emerging area of automation in management in an attempt to stimulate research on potential effects of automation on managers' job satisfaction and well-being at work.

Keywords Human-automation interaction · Automation in management · Artificial intelligence · Decision support systems · Personnel selection · Work characteristics

Introduction

For decades, automation has assisted human work (Sheridan & Parasuraman, 2005). Historically, automation predominantly affected production and aviation (Endsley, 2017) or supported monitoring tasks (e.g., monitor nuclear power plants; Muir, 1994). Nowadays, automation based on artificial intelligence and machine learning is more wide reaching and takes over more complex information processing and high-

inference information classification tasks assisting humans in many areas of everyday life. For instance, such highly automated information processing and decision support systems support judges at courts (Brennan, Dieterich, & Ehret, 2009), physicians with diagnoses (Bright et al., 2012), and high-stake managerial tasks (Langer, König, & Papathanasiou, 2019; Wesche & Sonderegger, 2019). Within human resource management, current practical developments and recent research point to a future where managers will collaborate with automated systems to work on tasks as varied as scheduling (Ötting & Maier, 2018), personnel selection (Campion, Campion, Campion, & Reider, 2016), and retention management (Sajjadiani, Sojourner, Kammeyer-Mueller, & Mykerezi, 2019).

Previous research regarding automation in management has focused mainly on questions surrounding efficiency and effectiveness. For instance, Campion et al. (2016) showed that automated processes might help to assess motivational letters, and Naim, Tanveer, Gildea, and Hoque (2018) showed that

automated processes can be used to evaluate job interviews. Another stream of research investigates if and how people use decision support by automated systems (Burton, Stein, & Jensen, 2019) and discusses whether people are averse to using recommendations generated by automated systems (Dietvorst, Simmons, & Massey, 2015) or if they appreciate such recommendations (Logg, Minson, & Moore, 2019).

However, implications of automation do not stop with matters of usage or questions surrounding efficiency and effectiveness. Classical areas of automation have shown that the implementation of automation affects work tasks, motivation, and in general, well-being at work (Smith & Carayon, 1995). We propose that automation supporting managerial tasks can have similar effects. Especially in personnel selection (as a specific managerial task), automation has already crucially affected the means of fulfilling a hiring manager's job. For instance, they might be supported by systems gathering applicant information and providing evaluations of candidates based on automatic screening of resumes (Campion et al., 2016) or job interviews (Langer et al., 2019). On the one hand, this potentially increases productivity and allows additional time to work on other tasks (Parasuraman & Manzey, 2010). On the other hand, automated systems affect hiring managers' information processing during hiring (e.g., reduce the need to process raw applicant data; Endsley, 2017; Parasuraman & Manzey, 2010) and they might perceive that their personal influence on the selection process is reduced (Nolan, Carter, & Dalal, 2016). Such effects may impact perceived autonomy and responsibility within the respective tasks which could affect work satisfaction and motivation (Hackman & Oldham, 1976; Humphrey, Nahrgang, & Morgeson, 2007).

In the current study, we bridge research on automation and reactions to information processing and decision support systems with propositions of work design research (Hackman & Oldham, 1976; Morgeson & Humphrey, 2006; Parker & Grote, 2020). Specifically, we propose that different versions of automated decision support systems can affect work characteristics. We investigate how different versions of those systems might affect human behavior and reactions while completing personnel selection tasks. Participants in the current study worked on five consecutive personnel selection tasks and either received no support, support from a system providing support (i.e., an automated ranking of applicants) *before* participants processed applicant information, or support provided by a system *after* initial human processing of applicant information (Endsley, 2017; Guerlain et al., 1999). In addition to examining the effectiveness and efficiency of working with the systems, we investigate participants' reactions to the task (e.g., enjoyment, self-efficacy), which are related to well-being at work (Chung-Yan, 2010; Morgeson, Garzsa, & Campion, 2012). It is our intent that this study will stimulate further research on the effects of automation in management.

Background and development of hypotheses

Automation and decision support for personnel selection

Automated support systems for high-level cognitive tasks are an emerging topic in research and practice, especially within personnel selection. For instance, Campion et al. (2016) used an automated system to emulate human ratings of applicant motivational letters. Such a system processes applicant information and could provide hiring managers with rankings of candidates to aid during hiring processes. Other studies have investigated the use of automated job interviews (Naim et al., 2018; Suen, Chen, & Lu, 2019). Those studies argue that it might be possible to automatically process interviews to predict interview performance which would make it feasible to screen thousands of applicants and present an evaluation of the most suitable applicants to hiring managers (Langer, König, & Hemsing, 2020).

The emerging use of such systems especially for personnel selection might be due to the complexity of personnel selection. Hiring managers gather and integrate a large variety of information (e.g., results from cognitive ability tests, interviews) from a potentially large number of applicants. Additionally, they screen the most suitable applicants, decide which applicants to hire, consider a variety of organizational goals (e.g., cost, diversity), and simultaneously need to adhere to legal regulations (e.g., regarding adverse impact) (König, Klehe, Berchtold, & Kleinmann, 2010). Automated decision support systems might help with gathering and combining of information for large numbers of applicants, thus making selection procedures more efficient, and there is hope that they can attenuate adverse impact (Campion et al., 2016; see, however, Raghavan, Barocas, Kleinberg, & Levy, 2020 showing challenges of automated personnel selection regarding adverse impact).

To get a clearer understanding of automated systems for personnel selection, we consider research in automation and on decision support systems. Following Sheridan and Parasuraman (2005, p. 89), automation refers to "a) the mechanization and integration of the sensing of environmental variables (by artificial sensors), b) data processing and decision making (by computers), c) mechanical action (by motors or devices that apply forces on the environment), d) and/or 'information action' by communication of processed information to people." We focus on the data processing and decision-making aspects of automation followed by information action where systems provide processed information to human decision-makers. This kind of automation is reflected by four broad classes of functions of automation: gathering of information, filtering and analysis of information, decision recommendations, and action implementation (Parasuraman, Sheridan, & Wickens, 2000; Sheridan, 2000). With more of

these functions realized within systems, the extent of automation increases. For instance, gathering of information could imply automatic transcription of video interviews (Langer, König, & Krause, 2017). Filtering and analysis of information could mean highlighting keywords within these transcripts. Systems that provide decision recommendations could be trained on previous data in order to fulfill classification (e.g., distinguishing suitable vs. non-suitable applicants) or prediction tasks (e.g., what applicants will most likely be high-performing employees). The outputs of those tasks can then be presented to human decision-makers. This is precisely the kind of automated system that the current paper is referring to: automated systems that gather and analyze information as well as derive inferences and recommendations based on these steps. Such systems can aid hiring managers in a large range of their duties (e.g., gathering and integrating of information, applicant screening), and their outputs can be used as an additional (or alternative) source of information for decision-makers.

Further research on automated information processing and decision support systems has its roots in the literature on mechanical versus clinical gathering and combination of information (Kuncel, Klieger, Connelly, & Ones, 2013; Meehl, 1954). In line with this research, automated processes from the era of machine learning and artificial intelligence can be perceived as a more complex and sophisticated way of mechanical gathering or combination of data. Specifically, some of the novel approaches use sensors (e.g., cameras, microphones) to extract information from interviews, use natural language processing to extract the content of applicant responses, and apply machine learning methods to evaluate applicant performance (how this novel kind of mechanical information gathering and combination compares to traditional mechanical approaches and to clinical approaches regarding validity is an open question; Gonzalez, Capman, Oswald, Theys, & Tomczak, 2019). Research on mechanical combination of information can provide insights into potential effects of using modern information processing and decision support automation (see, e.g., Gonzalez et al., 2019; Longoni, Bonezzi, & Morewedge, 2019). On the one hand, it has shown that mechanical gathering and combining of information (e.g., combination using ordinary least squares regression) can improve decision quality when compared to clinical gathering and combining of information (e.g., intuition-based combination of information) (Kuncel et al., 2013). On the other hand, this literature found that people are skeptical of the use of mechanical gathering and combining of information (Burton et al., 2019). Some researchers even refer to the reluctance to use mechanically combined information as algorithm aversion (Dietvorst et al., 2015; but cf. Logg et al., 2019) and indicate that using automated systems within jobs might affect people's behavior and reactions to the job (Burton et al., 2019).

Information processing and decision support systems and their effect on performance and efficiency

There are several ways in which automated information processing and decision support systems for personnel selection can be implemented. Beyond the specific tasks allocated to an automated system (e.g., information gathering, analysis of information), the timing of when to provide hiring managers with the outputs of a system is a crucial implementation choice (Silverman, 1992; van Dongen & van Maanen, 2013). Two of the main points in time to integrate decision support systems are systems that provide their outputs (e.g., rankings of applicants) before a human decision-maker processes available information (*support-before-processing* systems) and systems that provide their outputs after an initial processing of information (*support-after-processing* systems) (Endsley, 2017; Guerlain et al., 1999; Silverman, 1992). For instance, automated personnel selection systems can come with automated support-before-processing (see Raghavan et al., 2020 who provides an overview on providers of automated personnel selection solutions). In general, the respective system analyzes applicant data and derives outputs (e.g., evaluations of performance, personality; applicant rankings) (Langer et al., 2019). Hiring managers receive these outputs together with further applicant information. Thus, they receive the output of the automated system and can have an additional look into further applicant information. This means decision-makers could decide to fully rely on the recommendation provided by the system, but they can also use it as an additional source of information to integrate together with further applicant information to reach a decision (Kuncel, 2018). Potential advantages of these systems are that they can increase efficiency (Onnasch, Wickens, Li, & Manzey, 2014) and serve as a source of mechanically combined information to enhance decision quality (Kuncel et al., 2013)—given an adequate validation of the system. Potential disadvantages are that they can induce anchoring effects so hiring managers might only give attention to best-ranked applicants (Endsley, 2017). Additionally, research from classical areas of automation indicates that people tend to initially perceive such systems as highly reliable which can lead decision-makers to follow recommendations without consulting additional, potentially contradictory information, and without thoroughly reflecting on applicants' suitability (i.e., they might overtrust the system; Lee & See, 2004). Finally, for support-before-processing systems, people can feel "reduced to the role of [...] recipient of the machine's solution" (Silverman, 1992, p. 111). This feeling can diminish user acceptance and might be one reason for perceived loss of reputation when using such systems because decision-makers perceive they have less opportunity to show their expertise (Arkes, Shaffer, & Medow, 2007; Nolan et al., 2016).

Partly due to the latter issues, support-after-processing systems (Endsley, 2017; Guerlain et al., 1999) evolved as an

alternative to support-before-processing systems. They would also process applicant information and provide an evaluation of applicants. However, these systems serve as an additional source of mechanically combined information that decision-makers can use *after* they have processed available information (Silverman, 1992). Additionally, such systems can be designed to provide feedback or criticize human decisions (Sharit, 2003; Silverman, 1992). Rather than proving the correctness of human decision, those systems serve as an opportunity to reflect on an initial decision and as an additional point of view on the decision. Up to date, research on those systems comes primarily from medical research and practice (Longoni et al., 2019; Silverman, 1992). For instance, in cancer diagnosis, a physician would first analyze available data (e.g., MRI images; Langlotz et al., 2019) and come up with a diagnosis (or a therapy plan; Langlotz & Shortliffe, 1984). The respective systems could use this diagnosis as an input and either provide the physician with the diagnosis it would have given or with information regarding what part of a diagnosis seems to be inconsistent with available data (Guerlain et al., 1999). However, because there is already evidence that issues with support-before-processing systems (e.g., perceived loss of reputation) might translate to managerial tasks (Nolan et al., 2016), the use of support-after-processing systems will likely not remain restricted to medical decision-making. Optimally, they encourage more thorough information processing and finding better rationales for decisions (Guerlain et al., 1999). For instance, such systems in personnel selection might counterbalance human heuristics by making decision-makers aware of overlooked or hastily rejected candidates (Derous, Buijsrogge, Roulin, & Duyck, 2016; Raghavan et al., 2020). Further potential advantages are that those systems, similar to support-before-processing systems, serve as an additional mechanical source of information combination, thus potentially enhancing decision quality (Guerlain et al., 1999). Furthermore, and in contrast to support-before-processing systems, there are no initial anchoring issues when using support-after-processing systems (Endsley, 2017).

One disadvantage of support-after-processing systems is that they do not increase efficiency of decision-making as decision-makers still need to process initially available information. In contrast, they can even increase the time necessary to reach decisions as decision-makers are encouraged to consider additional information and alternative perspectives on available information (Endsley, 2017). Furthermore, over-trusting effects cannot be ruled out. However, instead of solely relying on the recommendation by an automated system (as could be the case for support-before-processing systems), decision-makers themselves would have processed and combined different sources of information. This could give them a better rationale whether or not they want to follow the system's recommendations (Sharit, 2003). Given the proposed advantages and disadvantages of the different decision

support systems, we propose (see Table 1 for an overview on contrasts for the hypotheses):

- Hypothesis 1[1]: Participants who receive support (by a support-before-processing system or a support-after-processing system) will show better performance in the tasks than participants who receive no support (i.e., the no-support group).
- Hypothesis 2[2]: Participants in the support-before-processing group will complete the decision in less time than the other groups.

Effects on knowledge and task characteristics

In a review on the topic of algorithm aversion, Burton et al. (2019) argue that using automated support systems might have a variety of effects on decision-makers. For instance, decision-makers might feel reduced autonomy when receiving decision support, expectations towards systems might not be met by respective systems (Dietvorst et al., 2015; Highhouse, 2008), and certain design options within such systems (e.g., how and when to present a recommendation) might not match to human information processing, thus contributing to reluctance to use such systems.

This implies that using such systems for managerial work may affect knowledge and task characteristics (e.g., during information processing and decision-making tasks), and we propose that work design research (Hackman & Oldham, 1976; Morgeson & Humphrey, 2006; Parker & Grote, 2020) can help to understand these implications. Specifically, the integrated work design framework (Morgeson et al., 2012) proposes a variety of knowledge and task characteristics that affect important attitudinal, behavioral, cognitive, and well-being outcomes. Knowledge characteristics refer to the demands (e.g., cognitive demands) that people experience while fulfilling tasks (Morgeson & Humphrey, 2006). Task characteristics relate to the tasks that have to be accomplished for a particular job and how people accomplish these tasks.

In relation to decision support systems, the knowledge characteristic information processing and the task characteristics autonomy, task identity, and feedback from the job are especially important. The amount of information processing reflects the degree to which a job requires processing, integration, and analysis of information (Morgeson &

---

[1] This study was preregistered under https://aspredicted.org/xk2tg.pdf. In the preregistration, we stated that only the support-before-processing group will perform better, and after reviewing the literature in more depth, theory clearly indicated otherwise.
[2] In the preregistration, we stated that the support-after-processing group will be faster than the no-support group, but after considering more of the literature, we could not support this assumption.

Table 1 Overview on contrasts for the hypotheses

	NS	SB	SA
Hypothesis 1 (performance)	−2	1	1
Hypothesis 2 (decision time)	1	−2	1
Hypothesis 3a (enjoyment)	1	−2	1
Monotony	−1	2	−1
Hypothesis 3b (enjoyment)	−1	–	1
Monotony	1	–	−1
Hypothesis 4 (satisfaction)	−1	−1	2
Hypothesis 5 (responsibility)	1	−2	1
Hypothesis 6a (specific SE) Hypothesis 6a (general SE)	Interaction contrast: increase from time 1 to time 5 larger for the support-after-processing group than for other groups		
Hypothesis 6b (specific SE) Hypothesis 6b (general SE)	Dependent *t* test: increase from time 1 to time 5 only for the support-after-processing group and the no-support group		

The numbers in the columns NS, SB, and SA reflect the contrast codes for the hypotheses. A coding of 2 or − 2 indicates that this group was compared to the mean of the groups coded with − 1 or 1 (because contrast weights have to sum up to zero). The positive numbers for the codes indicate that the mean values for the respective groups for the respective dependent variable were expected to be higher. A dash (–) indicates that this group was ignored for the respective comparison.

NS no-support group, *SB* support-before-processing group, *SA* support-after-processing group, *SE* self-efficacy

Humphrey, 2006). Autonomy indicates the degree to which employees can fulfill tasks independently and can decide how to approach tasks (Hackman & Oldham, 1976; Morgeson et al., 2012). Task identity describes the degree to which tasks can be fulfilled from the beginning to the end versus only working on specific parts of the task (Hackman & Oldham, 1976; Morgeson et al., 2012). Finally, feedback from the job is defined as the degree to which employees receive information about their performance in the job from aspects of the job itself (Hackman & Oldham, 1976; Morgeson et al., 2012). Variations in these characteristics affect psychological states such as experienced meaningfulness, perceived responsibility for work outcomes, as well as work satisfaction and performance (Chung-Yan, 2010; Hackman & Oldham, 1976; Humphrey et al., 2007). While Humphrey et al. (2007) emphasize the lack of empirical research regarding knowledge characteristics, previous research shows that, in general, more demanding information processing requirements, a higher level of autonomy, task identity, and feedback from the job relate to more positive outcomes (Morgeson et al., 2012; but see Chung-Yan (2010) and Warr (1994), indicating that "the more the better" is not necessarily true for all characteristics).

Information processing and decision support systems for personnel selection might affect those knowledge and task characteristics. Support-before-processing systems show their assessment of applicants before hiring managers process applicant data. This might reduce information processing requirements (e.g., integrate information, compare applicants). Regarding task identity, hiring managers might feel that they do not really complete the entire selection task. Moreover, they might perceive a loss of autonomy

as the system already implies which applicants to favor (Burton et al., 2019; Langlotz & Shortliffe, 1984; Nolan et al., 2016). All of this would be in line with findings and speculations of previous research that indicated that hiring managers tend to favor selection methods where they can show their expertise and those that allow for more autonomy (e.g., unstructured vs. structured interviews; using clinical vs. mechanical combination of information) (Burton et al., 2019; Highhouse, 2008).

In contrast, support-after-processing systems might require more information processing, allow for more perceived task identity, and grant a higher level of autonomy. They allow hiring managers to independently analyze and integrate information, reach an initial decision about applicants, and only then provide them with additional information or novel perspectives for their final decision. Furthermore, hiring managers might use recommendations from support-after-processing systems as feedback on task performance. Specifically, when hiring managers have no prior experience with an automated system, they may believe that the system is working as intended (Madhavan & Wiegmann, 2007). The purpose of an automated personnel selection system is to evaluate applicants' job fit, so if people initially believe that an automated system is able to do this, they might compare their own decisions to the recommendations by the system in order to get an idea about their own performance.

Following, we will build hypotheses that argue for how the aforementioned variations in knowledge and task characteristics affect five important psychological factors at work: enjoyment, monotony, satisfaction with the decision, perceived responsibility, and self-efficacy.

Effects on enjoyment and monotony

Enjoyment and monotony of a task are important for well-being at work (Taber & Alliger, 1995). Enjoyment of a task is present if people are happy doing a task, whereas monotony indicates that people perceive a task to be repetitive and boring (Smith, 1981). The proposed differences in information processing requirements, task identity, and perceptions of feedback might affect enjoyment and experienced monotony with the task (Smith, 1981). In the case of a support-before-processing system, the task might appear less cognitively demanding, and decision-makers might think that most of the task was already fulfilled by the system. In contrast, using support-after-processing systems upholds information processing requirements, and task identity. Additionally, if people perceive that they are provided with evidence on how well they performed, this can increase enjoyment within a task (Sansone, 1989). Thus, we propose:

- Hypothesis 3a: Participants in the support-before-processing system group will perceive less enjoyment and more monotony with the task than the other groups.
- Hypothesis 3b: Participants in the support-after-processing system group will perceive more enjoyment and less monotony with the task than the no-support group.

Effects on satisfaction with the decision

For knowledge workers (e.g., hiring managers), a large share of daily work consists of information processing and decision-making (Derous et al., 2016). Satisfaction with decisions thus likely contributes to workers' overall job satisfaction (Taber & Alliger, 1995). Satisfaction with decisions is especially important in situations where the consequences and the quality of a decision do not immediately become apparent (Houston, Sherman, & Baker, 1991; Sainfort & Booske, 2000). This is the case for personnel selection, where the quality of the decision is determined by an applicant's future job performance (Robertson & Smith, 2001). When the consequences of decisions will only become visible in the long term, satisfaction with decisions might arise if people are convinced by their decision and satisfied with the decision-making process (Brehaut et al., 2003; Sainfort & Booske, 2000). Especially when people autonomously process information, they might become more convinced by their decision and thus more satisfied with it (Sainfort & Booske, 2000). Additionally, if people believe they have made a good decision (i.e., if they get any kind of perceived feedback on their performance), this might increase satisfaction with the decision (Sansone, 1989). To be clear, information processing requirements, task identity, task autonomy, and feedback might be more

pronounced in the case of support-after-processing systems; thus, we propose:

- Hypothesis 4[3]: Participants in the support-after-processing system group will be more satisfied with their decision than the other groups.

Effects on perceived responsibility

Research from different areas of decision support shows that using standardized or computerized processes can reduce users' perceived responsibility (Arkes et al., 2007; Lowe, Reckers, & Whitecotton, 2002; Nolan et al., 2016). According to Nolan et al. (2016), perceived responsibility in personnel selection may decrease because decision-makers believe that task fulfillment cannot be fully attributed to themselves. Instead, some of the credit goes to the automated or standardized support process. Aligning with this, work design research assumes that reduced autonomy leads to less perceived responsibility (Hackman & Oldham, 1976). Furthermore, Langer et al. (2019) speculated about situations where hiring managers who are supported by a system follow the advice of the system. Hiring managers might perceive that the decision was already predetermined, so even if they were to thoroughly evaluate all available information, by following the system's recommendation, their perceived responsibility could reduce, especially for those using a support-before-processing system. This is why we propose:

- Hypothesis 5: Participants in the support-before-processing system group will feel less responsible for their decision than the other groups.

Effects on self-efficacy

Self-efficacy reflects a person's belief that through their skills and capabilities, they can show a certain task performance (Bandura, 1977; Stajkovic & Luthans, 1998). Self-efficacy is an important work-related variable as it affects job performance and work satisfaction (Judge, Jackson, Shaw, Scott, & Rich, 2007; Lent, Brown, & Larkin, 1987; Stajkovic & Luthans, 1998).

Self-efficacy can be task specific or general (Jerusalem & Schwarzer, 1992; Schwarzer & Jerusalem, 1995). Specific self-efficacy strongly depends on a specific situation (Trippel, 2012), whereas general self-efficacy would generalize to a wider variety of situations (Jerusalem & Schwarzer,

---

[3] In the preregistration, we suggested that satisfaction in the support-before-processing group would be lower; however, the theoretical rationale for this claim was weaker than that for what now constitutes Hypothesis 4.

1992). In the current study, specific self-efficacy relates to the perceived self-efficacy to perform the selection task at hand, whereas general self-efficacy relates to participants' self-efficacy in personnel selection as a whole.

Successfully experiencing and independently fulfilling a task should strengthen both specific and general self-efficacy (Bandura, 1977). This should be especially pronounced for more demanding tasks (e.g., tasks that afford more information processing). Furthermore, receiving evidence of good performance can increase self-efficacy (Bandura, 1977). Describing the differences between the experimental groups in the terminology of work design research, only in the support-after-processing system and in the no-support group, participants may perceive that they autonomously fulfilled an entire cognitively demanding task. Additionally, participants in the support-after-processing condition might interpret the information by the system as an indicator of their task performance. All of this could strengthen specific and general self-efficacy.

- Hypothesis 6a: Participants in the support-after-processing system group will show a higher increase in the degree of general and specific self-efficacy than the other groups.

In contrast, participants supported by the support-before-processing system might believe that they did not fulfill the entire task themselves, might perceive the task to be less demanding, and might experience less autonomy (Burton et al., 2019). Those participants could also believe that they were only following the advice of the system (Endsley, 2017; Langer et al., 2019). If this is the case, they might not perceive that they contributed to the task beyond what was already given by the system so completing the task would neither strengthen their feelings of being capable of performing the task at hand nor would it translate to self-efficacy in personnel selection in general.

- Hypothesis 6b: General and specific self-efficacy will only increase for participants in the support-after-processing group and for the no-support group.

Method

Sample

We consulted work design research in order to estimate expectable effect sizes between our experimental conditions. Humphrey et al. (2007) showed relations between different psychological states (e.g., responsibility) and autonomy ($r = .27$–$.51$), task identity ($r = .20$–$.27$), and feedback ($r = .33$–$.46$). Wegman, Hoffman, Carter, Twenge, and Guenole (2018) reported effect sizes in a similar range for task characteristics and job satisfaction (autonomy, $r = .39$; task identity, $r = .26$; feedback, $r = .33$). Since there was a lack of informative studies regarding performance and efficiency gains from the use of automation in managerial tasks, we consulted the literature from different domains. Bright et al. (2012) reported performance gains in clinical decision-making of $d = 0.30$. Crossland, Wynne, and Perkins (1995) showed that using a decision support system in a complex decision-making task can reduce necessary time to decide by $d = 0.58$. Overall, we thus assumed a medium effect size in the current study. This might seem rather progressive in the case of performance; however, our decision support system was perfectly reliable, so we deemed a medium effect size expectable. We determined the required sample size with G*Power (Faul, Erdfelder, Buchner, & Lang, 2009). Assuming a medium between-conditions effect of $\eta^2_p = 0.06$ with three groups in an ANOVA, $N = 117$[4] participants would be necessary to achieve a power of $1 - \beta = .80$. This study was advertised to people interested in human resource management. We posted the advertisement on different social media groups, around the campus of a medium-sized German university, and in the inner city of a medium-sized German city.

As we anticipated potential issues during data collection (e.g., participants showing insufficient effort), we continued data collection until our sample consisted of $N = 131$ participants. We excluded four participants because they did not finish the experiment, three because they reported that they did not understand the task, one participant for showing insufficient effort (responding "neither agree nor disagree" to all items), and one participant because of technical issues. The final sample consisted of $N = 122$ German participants (71% female), 71% students (of which 88% studied psychology or business, with the rest of students coming from a variety of disciplines). Of the remaining participants, more than half were employed full-time coming from a variety of backgrounds. The mean age was 25.52 years (SD = 9.31), and participants had experienced a median of three personnel selection processes as applicants in their lives. Psychology students were rewarded with course credit, and all participants could win vouchers for online shopping.

Procedure

Figure 1 provides an overview of the study procedure and measures. Participants came to the lab and were instructed to sit in front of a computer where they were guided through the study using a survey tool. Participants were randomly assigned to one of three conditions (no-support, support-

---

[4] In the preregistration, we mention an N value of 111 participants which was based on an incorrect calculation.

Fig. 1 Study procedure and measures. The left stream in the figure depicts the procedure for the support-before-processing group, the middle stream depicts the procedure for the no-support group, and the right stream for the support-after-processing group. DV dependent variables. Note that DVs2 and DVs3 are the second and third measures, respectively, of the dependent variables enjoyment, monotony, satisfaction, perceived responsibility, and specific self-efficacy in the task
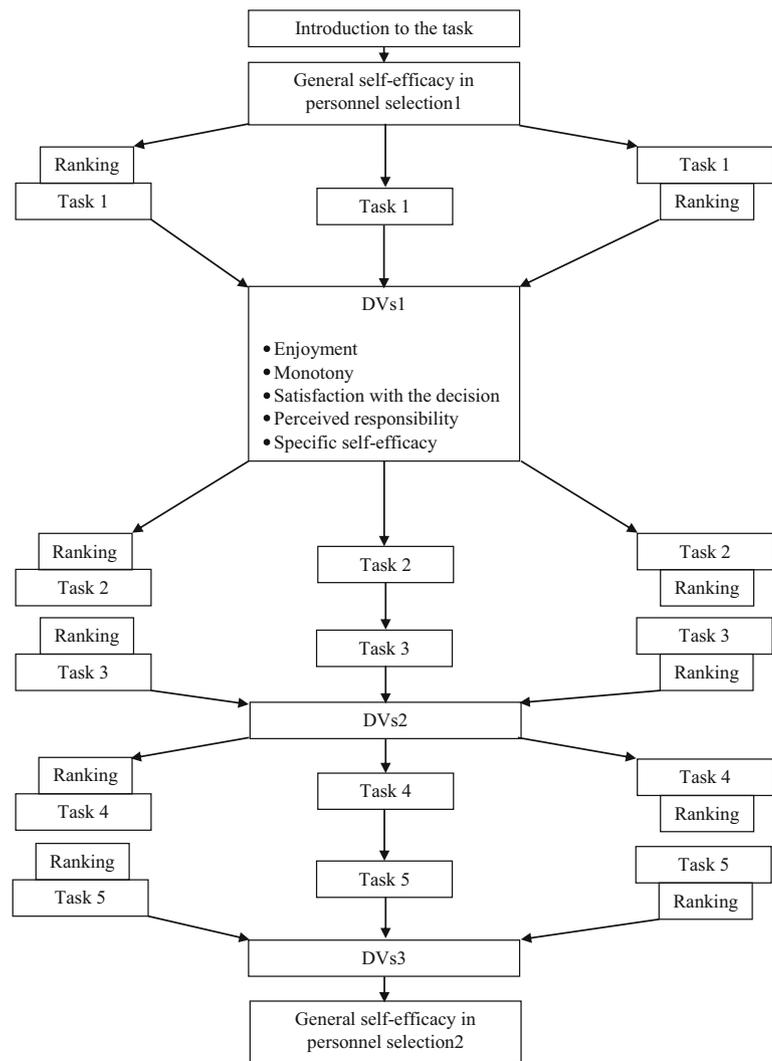
before-processing, support-after-processing). They were told to imagine themselves as hiring managers for a telecommunication company searching for sales managers for different subsidiaries. Participants were provided with a printed job description (adapted from actual job descriptions of sales managers in the telecommunication industry) describing the necessary knowledge, skills, abilities, and other characteristics (KSAOs) of sales managers (e.g., experience in sales, customer orientation).

After reading the job description, participants experienced five hiring tasks. For each task, participants had to decide, out of six applicants, who they would hire. For each applicant, participants received asynchronous interview-like self-presentation audio files (Langer et al., 2020) (see section "Development of the support systems" for further information) that participants could listen to in order to decide whom to hire.

In the no-support condition, participants only received applicants' self-presentations and then had to decide which applicant they would hire in a given task round. In both conditions that received automated support, participants were informed that an automated system had analyzed and evaluated applicants' self-presentation audio files and generated a ranking based on the estimated job fit of the respective applicant. The ranking categorized applicants into A (high job fit), B (medium job fit), or C (low job fit) applicants (two applicants per category). Note, that we conducted a pre-study (see section "Development of the support systems") where we determined that the self-presentation audio files reflected applicants' level of job fit. This means the systems' rankings reliably distinguished applicants into the three categories of job fit. Participants received no information regarding the systems' reliability. In the support-before-processing condition, participants received the ranking before they accessed applicants' self-presentations. In the support-after-processing condition, participants received the same ranking after their initial processing of information (indicated by participants making an initial decision for one of the applicants). Those

participants then had the opportunity to use the ranking as an additional source of information, to reflect on their decision and received the opportunity to potentially change their initial decision which is common within support-after-processing systems (Endsley, 2017; Guerlain et al., 1999).

Development of the support systems

Based on the job description, we developed profiles of six applicants for all five selection tasks. Specifically, we analyzed the job description regarding KSAOs, made a list of the required KSAOs, and developed applicant profiles that fit the required KSAOs well (A applicants), showing a medium fit (B applicants), and showing bad fit (C applicants). Based on these profiles, we developed 30 texts (5 hiring tasks, 6 applicants for each task) similar in the number of words where the supposed applicants introduce themselves (i.e., past experience, education, skills). For instance, A applicants indicated more information that suited the job description (e.g., they talked about past jobs where they gained sales experience). In contrast, C applicants talked about their prior experience in jobs unrelated to sales or were underqualified. To examine the distinguishability of the introductory texts regarding suitability for the job, we tested them in a pre-study. Participants ($N = 25$) received the job description and read the self-presentation texts of ten of the applicants in a random order. For each applicant, participants responded to the item "This applicant is suited for the position" on a scale from 1 (not at all) to 7 (very good). Participants in the pre-study were able to distinguish between A, B, and C applicants. Note that this way, we also ensured the reliability of the support systems. This means, with the pre-study, we made sure that the rankings provided by the systems reliably reflect applicants' quality based on our categorization into A, B, or C applicants.

We then recruited 30 speakers with an age between 25 and 35 years to learn the texts and record an audio version of those texts. The speakers were instructed to imagine being in a personnel selection situation where they were asked to present themselves, their experience, skills, and abilities, to a hiring company. The texts were written in a way to keep the duration of the audio recordings at about 1 min in order not to overtax participants since every participant received a total of 30 audio recordings. The audio recordings should be similar to asynchronous interview recordings where applicants present themselves to an organization (see Ambady, Krabbenhoft, and Hogan (2006), showing that very short interactions can be enough to predict performance of sales people, or see Hiemstra, Oostrom, Derous, Serlie, and Born (2019) or Langer et al. (2020) where participants in an interview study recorded asynchronous audio files of between 1 and 1.5 min). This was done in order to enhance realism of the hiring context and to require participants to open audio files if they want to access information about the applicants.

Measures

Participants responded to all self-report measures on a scale from 1 (strongly disagree) to 7 (strongly agree). All measures were administered in German. One of the authors translated the self-report measures from English, another author translated them back to English to check for consistency, and finally, potential inconsistencies were discussed and resolved.

Enjoyment and monotony were measured with six items each, taken from Plath and Richter (1984). A sample item for enjoyment is "Right now, I have fun with this task"; for monotony, "Right now, the task makes me bored."

Satisfaction with the decision was measured with four items of the decision regret scale (Brehaut et al., 2003). A sample item is "This was the right decision."

Perceived responsibility for the decision was measured with three items from Nolan et al. (2016). We adapted these items so participants could answer them in a first-person perspective. A sample item is "I feel responsible for the outcome of this hiring decision."

General self-efficacy in personnel selection was measured with eight items from the scale professional self-efficacy expectations by Schyns and Von Collani (2014) that were adapted to fit the context of personnel selection. A sample item is "I feel confident regarding most of the demands in personnel selection." Participants were instructed to evaluate their own beliefs about their abilities to fulfill personnel selection tasks in general.

Specific self-efficacy in the task was measured with six items. Of these items, three were adapted from Trippel (2012) who developed the items based on Luszczynska, Scholz, and Schwarzer (2005). A sample item is "I am convinced that I can be successful in this task." In order to capture if participants evaluate their specific self-efficacy regarding the task differently when explicitly mentioning help that they might receive (i.e., assuming that people might perceive the automated system as "help"), we added the phrase "without help" to each of the items and added those items as three more items. A sample item is "I am convinced that I can be successful in this task without help." Note that there were no differences in the results when analyzing a scale based on the first versus the second three items of this scale which is why report results for all six items combined into a scale.

Performance. Participants received 1 point for each time they decided for one of the A candidates, 0.5 point when they decided for one of the B candidates, and 0 point otherwise.

Time to decide was captured through the time-log functionality of the questionnaire tool. Specifically, we calculated the time between the start of a task round (i.e., when participants landed on the instruction page for the respective task round) and the end of a task round (i.e., the final decision for an applicant).

Results

Analyses were conducted with R, version 3.6.0. Table 2 provides correlations, means, and *SD*s for all variables over all tasks. In order to examine if dependent variables could be distinguished, we examined our measurement model using confirmatory factor analyses including the dependent variables after each task round. We compared our measurement model (the dependent variables load on a separate latent factor, and those factors are correlated) to two alternative models (one-factor model; enjoyment and monotony load on a common factor). Results indicated that our measurement model (ranges for all tasks, CFI = .86, RMSEA = .09–.10, SRMR = .07) showed a slightly better fit than the other models (ranges for all tasks, one-factor model: CFI = 83–.85, RMSEA = .09–.11, SRMR = .09–.13; enjoyment and monotony on a common factor: CFI = .82–.86, RMSEA = .10, SRMR = .07–.08) (see Table 3). The overall fit of our model after including theoretically sound modifications (e.g., including correlated residuals for items loading on one factor) was acceptable (for all tasks, CFI = .88–.91, RMSEA = .07–.09, SRMR = .07).

Figure 2 shows line graphs for all dependent variables, and Tables 4 and 5 show results of mixed ANOVAs. Within-participant effects over time were significant for all dependent variables except performance (see "Hypotheses testing"). This means, time to decide, enjoyment and satisfaction decreased, monotony, and responsibility increased. Specific and general self-efficacy also increased, but this was qualified by interaction effects (see "Hypotheses testing"). For all tasks combined, enjoyment and monotony showed high negative correlations. Additionally, participants who enjoyed the task more and who perceived it to be less monotonous took more time to decide. Furthermore, participants who felt more responsible and were more satisfied with their decision also felt higher self-efficacy.

Hypotheses testing[5]

Table 6 provides an overview on contrasts used for the analyses of hypotheses and their results. Hypothesis 1 stated that participants who receive support will show better performance in the task than the no-support group. Starting from task 4 (see Fig. 2 and Table 7), there was a lack of variance in performance evaluations as most participants chose an A applicant. We therefore only analyzed tasks 1–3 for Hypothesis 1. Additionally, there were only 4 cases overall where

participants decided for one of the C applicants. We therefore dichotomized the outcome (i.e., choosing an A applicant or a B applicant). For the analysis of Hypothesis 1, a repeated measures logistic regression including a random intercept for participants revealed that the effect between the no-support and the other groups approached significance Wald $\chi^2(1) = 3.24, p = .07$. There was no significant effect for the task round (Wald $\chi^2(2) = 1.00, p = .60$) or the interaction of the group and the task round (Wald $\chi^2(2) = 3.37, p = .19$). As an alternative analysis strategy, we calculated Fisher's exact tests for each of the first three rounds (for these analyses, we kept cases where participants chose a C applicant). Table 7 shows that the differences between the groups approached significance only for task 1. Overall, Hypothesis 1 was therefore not supported.

Hypothesis 2 proposed that participants in the support-before-processing group will complete the decision in less time than the other groups. Neither the main effect for the groups nor a specific contrast testing this hypothesis (see contrast results in Table 6) was significant; therefore, Hypothesis 2 was not supported.

Hypothesis 3a stated that participants in the support-before-processing group will perceive less enjoyment and more monotony with the task than the other groups. There was no main effect for enjoyment and monotony, and specific contrast testing Hypothesis 3a showed no significant differences; thus, Hypothesis 3a was not supported.

Hypothesis 3b stated that participants in the support-after-processing group will perceive more enjoyment and less monotony with the task than the no-support group. Contrasts results showed that the support-after-processing group enjoyed the task slightly more than the no-support group, but there was no significant difference for monotony. Thus, Hypothesis 3b was only partially supported.

Hypothesis 4 stated that participants in the support-after-processing group will be more satisfied with their decision than the other groups. There was a significant main effect between the groups. A contrast testing the hypothesis revealed that the support-after-processing group was more satisfied with the decision. Therefore, Hypothesis 4 was supported.

Hypothesis 5 stated that participants in the support-before-processing group will feel less responsible for their decision than the other groups. Neither the main effect for the groups nor a contrast testing this hypothesis was significant; therefore, Hypothesis 5 was not supported.

Hypothesis 6a proposed that participants in the support-after-processing group will show a higher increase in the degree of general and specific self-efficacy than the other groups. For both specific and general self-efficacy, there was an interaction effect between the task round and the groups. Interaction contrasts revealed that general self-efficacy and specific self-efficacy increased more strongly for the support-after-processing group than for other groups. This supports Hypothesis 6a.

---

[5] We added participants' prior experience with application processes (as applicants) and the variable students versus other participants as control variables to the analyses. This only led to numeric changes in results that did not affect interpretations of the results regarding hypotheses. Additionally, we analyzed all dependent variables with hierarchical linear models. The results did not change in a way that would have led to different interpretations.

Table 2 Means, standard deviations, and correlations of the mean of the dependent variables over all decision tasks

Variable	M	SD	1	2	3	4	5	6	7
1. Performance	0.94	0.09							
2. Time to decide	352.60	75.71	− .15						
3. Enjoyment	5.74	0.79	.02	.19*					
4. Monotony	2.87	0.96	− .02	− .27**	− .81**				
5. Satisfaction	5.83	0.55	.10	.03	.45**	− .35**			
6. Responsibility	5.72	0.77	− .00	.05	.39**	− .35**	.47**		
7. Specific SE	5.37	0.89	− .09	.04	.42**	− .32**	.56**	.63**	
8. General SE	4.82	0.89	− .12	− .01	.47**	− .36**	.48**	.43**	.70**

Time to decide is expressed in seconds. $N = 122$

SE self-efficacy

**$p < .01$; *$p < .05$

Hypothesis 6b stated that general and specific self-efficacy will only increase for participants in the support-after-processing group and for the no-support group. For Hypothesis 6b, we calculated dependent t tests for the difference between tasks 1 and 5 for all groups. Results indicated that general self-efficacy increased for all groups. Specific self-efficacy increased for all but the no-support groups. Thus, Hypothesis 6b was not supported.

Discussion

With the introduction of automated systems based on artificial intelligence and machine learning, managerial jobs change. The goal of this study was to investigate how different versions of automated information processing and decision support systems might affect reactions to fulfilling personnel selection as a specific managerial task. Based on work design research (Hackman & Oldham, 1976; Morgeson et al., 2012, Parker & Grote, 2020)

Table 3 Model fit indices for the measurement model, two alternative models, and after modifications

Model	χ^2 (df)	$\Delta\chi^2$ (df)	AIC	CFI	RMSEA	SRMR
Task 1						
1. Hypothesized correlated 5-factor model	570.10** (289)	–	7355.2	.86	.09	.07
2. 1-factor model	605.39** (294)	35.29** (5)	7381.0	.85	.09	.09
3. Enjoyment + monotony as 1 factor	664.37** (293)	82.26** (4)	7496.5	.82	.10	.07
After modification	460.52** (282)	− 109.58** (7)	7260.1	.91	.07	.07
Task 3						
1. Hypothesized correlated 5-factor model	618.13** (289)	–	7513.7	.86	.10	.07
2. 1-factor model	704.11** (294)	85.97** (5)	7589.7	.83	.11	.13
3. Enjoyment + monotony as 1 factor	674.18** (293)	56.04** (4)	7561.8	.84	.10	.08
After modification	562.29** (282)	− 55.84** (7)	7471.9	.88	.09	.07
Task 5						
1. Hypothesized correlated 5-factor model	627.80** (289)	–	7577.8	.86	.10	.07
2. 1-factor model	735.88** (294)	108.08** (5)	7725.5	.83	.11	.13
3. Enjoyment + monotony as 1 factor	673.94** (293)	46.15** (4)	7616.0	.86	.10	.07
After modification	551.06** (282)	− 76.74** (7)	7515.1	.90	.09	.07

$\Delta\chi^2$ indicates the difference between model 1 and the respective model. $N = 122$

AIC Akaike information criterion, CFI comparative fit index, RMSEA root mean square error of approximation, SRMR standardized root mean square residual

**$p < .01$

Fig. 2 Graphs for the mean values of the dependent variables throughout the tasks. T task, SE self-efficacy. Error bars indicate standard errors. Note that enjoyment, monotony, satisfaction with the decision, perceived responsibility, and specific self-efficacy were measured after participants' decision for tasks 1, 3, and 5. Performance and time to decision were measured for all tasks. General self-efficacy was measured before the tasks and after the tasks
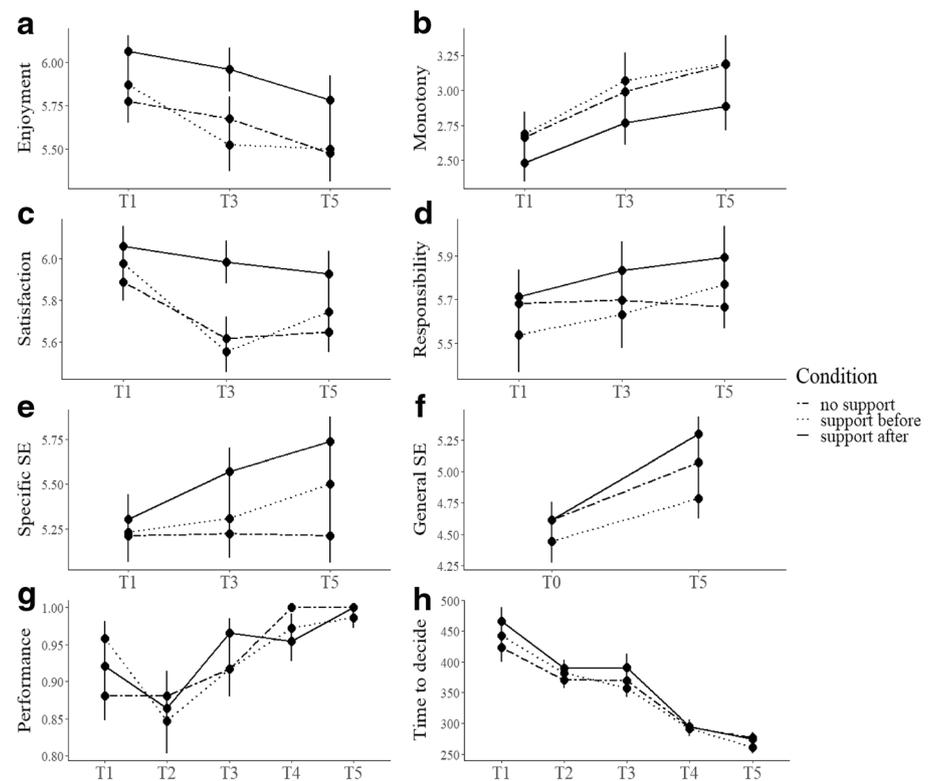
and research regarding reactions to decision support systems (Burton et al., 2019), we proposed that support systems will change knowledge and task characteristics influencing important psychological outcomes that relate to well-being at work (Morgeson et al., 2012). The first implication of the current study is that different versions of those systems can indeed affect important psychological outcomes. Specifically, support-after-processing systems more positively affected participants' satisfaction with decisions and self-efficacy. The second main implication is that we found no significant effectiveness and efficiency benefits from the use of the systems. However, we have to emphasize to draw conclusions based on this finding cautiously because participants quickly learned the task and hit the performance ceiling in a simulated personnel selection task without any time pressure.

Theoretical implications

Regarding task behavior (i.e., performance, time to reach the decision), only for the first task round there was a small but non-significant effect of receiving decision support regarding performance and there were no efficiency benefits in the groups that received decision support. This stands in contrast to previous finding from the area of mechanical gathering and combination of information (Kuncel et al., 2013) and to speculations about higher effectiveness and efficiency when using decision support systems (Burton et al., 2019; Endsley, 2017). Note, however, that the tasks seemed to be rather easy to learn

for participants. Specifically, performance improved for all groups to nearly perfect performance for the last two tasks, thus leading to reduced variance between the groups.

Regarding time to reach the decision, one could have expected faster decisions for the support-before-processing group—at least after experiencing that the system is reliable. This was not the case, potentially because participants still wanted to verify the recommendation by the unfamiliar system where participants had no information about the reliability of the system. Note, however, that in line with what can be expected (Endsley, 2017), descriptive results indicate that the support-after-processing group took more time to decide. Additionally, the no-support group did not have to process the additional information from the support systems which is reflected in the mean values (see Table 4), implying that this group completed the first two tasks in less time. However, mean values in the case of the support-before-processing group show that this group completed the last three tasks in less time, tentatively indicating that efficiency benefits from using support-before-processing systems develop when decision-makers are more familiar with the system.

The non-significant findings for task performance and time to make the decision, together with the finding that some participants disagreed with the system in task 2 (indicated by the decline in performance for task 2), could indicate fruitful directions for future research. Specifically, future research should examine effects on task behavior when tasks are harder and when automated decision support systems are less than

Table 4 Results of the mixed ANOVA for the dependent variables time to decide, enjoyment with the task, monotony with the task, satisfaction with the decision, responsibility for the decision, and specific self-efficacy

Variable	T1			T2			T3			T4			T5			ANOVA Condition		Within-effect		Interaction	
	NS	SB	SA	NS	SB	SA	NS	SB	SA	NS	SB	SA	NS	SB	SA	$F_{(1,119)}$	η^2_p	$F_{(1,119)}$	η^2_p	$F_{(2,119)}$	η^2_p
	M (SD)	M (SD)	M (SD)	M (SD)	M (SD)	M (SD)	M (SD)	M (SD)	M (SD)	M (SD)	M (SD)	M (SD)	M (SD)	M (SD)	M (SD)						
Time to decide	423 (154)	442 (138)	466 (148)	370 (85)	382 (118)	390 (93)	369 (93)	357 (94)	391 (146)	294 (73)	291 (75)	295 (75)	277 (59)	260 (57)	274 (50)	0.64	0.01	97.53**	0.45	0.81	0.01
Enjoyment	5.77 (0.80)	5.87 (0.76)	6.06 (0.60)	—	—	—	5.68 (0.84)	5.52 (0.92)	5.95 (0.84)	—	—	—	5.47 (1.04)	5.50 (0.99)	5.78 (0.93)	2.08	0.03	14.27**	0.11	0.93	0.02
Monotony	2.66 (0.90)	2.68 (0.96)	2.47 (0.87)	—	—	—	2.99 (1.01)	3.07 (1.17)	2.77 (1.07)	—	—	—	3.19 (1.09)	3.19 (1.23)	2.89 (1.16)	0.99	0.02	26.35**	0.18	0.20	0.00
Satisfaction	5.89 (0.58)	5.98 (0.63)	6.06 (0.62)	—	—	—	5.61 (0.67)	5.56 (0.60)	5.98 (0.69)	—	—	—	5.65 (0.64)	5.74 (0.73)	5.93 (0.74)	3.15*	0.05	10.50**	0.08	1.51	0.03
Responsible	5.68 (0.65)	5.54 (1.01)	5.71 (0.82)	—	—	—	5.70 (0.70)	5.62 (0.91)	5.83 (0.88)	—	—	—	5.67 (0.65)	5.77 (0.87)	5.89 (0.94)	0.53	0.01	3.98*	0.03	1.37	0.02
Specific SE	5.21 (0.95)	5.23 (0.99)	5.30 (0.93)	—	—	—	5.22 (0.87)	5.31 (0.98)	5.57 (0.87)	—	—	—	5.21 (0.96)	5.50 (0.99)	5.74 (0.91)	1.46	0.02	11.42**	0.09	3.79**	0.06

Time to decide is expressed in seconds. n_{NS} = 42, n_{SB} = 36, n_{SA} = 44

NS no-supportgroup, SB support-before-processing group, SA support-after-processing group, SE self-efficacy

**$p < .01$; *$p < .05$

perfectly reliable. This is important as less than perfectly reliable (in the sense of capable of providing good decision support) is what users can expect when using decision support systems for management tasks in practice. In the study by Campion et al. (2016) who investigated automatic scoring of motivational letters, correlations between the human raters and the automatic predictions were between .59 and .68. Naim et al. (2018) investigated the automatic analysis of job interviews and showed correlation coefficients of .65 for overall hiring recommendations between automatically generated ratings and ratings provided by MTurk workers. There is still a lack of validation effort for such systems (Raghavan et al., 2020), but it is possible that they will not demonstrate much higher validity. Considering the scarce evidence regarding validity of automatic systems for management tasks, receiving decision support could even result in lower decision quality than completing tasks without support (see, e.g., Onnasch et al., 2014, for similar results in classical areas of automation). For instance, in the case of support-before-processing, managers might blindly follow the systems' recommendation (Parasuraman & Manzey, 2010). In the case of support-after-processing, less-than-perfect decision support might lead to confusion and longer time spent on the decision and may lead hiring managers to rethink good decisions (Endsley, 2017).

Systems' reliability is a determining factor that affects reactions to automated systems (Hoff & Bashir, 2015; McKnight, Carter, Thatcher, & Clay, 2011). Similar to how trust in automated systems can decline drastically when perceiving errors by automated systems (de Visser, Pak, & Shaw, 2018; Madhavan & Wiegmann, 2007), people's perceptions of tasks supported by automated system can change drastically when people experience system failures. This brings us to the second main topic of the current study, where we respond to calls for research investigating the effects of technology on work design outcomes (especially knowledge characteristics such as information processing; Humphrey et al., 2007; Oldham & Fried, 2016, Parker & Grote, 2020). Our study indicates that similar to classical areas of automation (e.g., production, aviation; Endsley, 2017), automation in management can affect psychological reactions. We argue that this might be due to their influence on knowledge and task characteristics (Morgeson et al., 2012; Parker & Grote, 2020).

Specifically, the support-before-processing system in the current study supported participants and pointed them reliably towards the best applicants. However, participants were comparably less satisfied with their decisions. In contrast, the support-after-processing system might have been perceived as upholding autonomy and, in most cases, confirmed participants' initial decisions (in 84% of cases). As a result, participants were more satisfied with their decisions and more strongly increased their self-efficacy. However, what would

Table 5 Results of the mixed ANOVA for the dependent variable general self-efficacy

Variable	Before the tasks			After the tasks			ANOVA					
	NS	SB	SA	NS	SB	SA	Condition		Within-effect		Interaction	
	M (SD)	M (SD)	M (SD)	M (SD)	M (SD)	M (SD)	$F_{(1, 119)}$	η^2_p	$F_{(1, 119)}$	η^2_p	$F_{(2, 119)}$	η^2_p
General self-efficacy	4.62 (0.91)	4.44 (1.01)	4.61 (0.92)	5.07 (0.90)	4.78 (0.96)	5.30 (0.90)	1.50	0.03	97.44**	0.45	4.07*	0.06

$n_{NS} = 42$, $n_{SB} = 36$, $n_{SA} = 44$

NS no-support group, SB support-before-processing group, SA support-after-processing group.

**$p < .01$; *$p < .05$

have happened if the system had been less reliable? Potentially, participants would have been confused by the recommendation, and reactions to the task would have been similar to those for support-before-processing systems. This could be tested by future research varying the reliability of decision support systems (see Dietvorst et al., 2015 and Logg et al., 2019 for examples of such studies in other areas). This could clarify why our participants preferred support-after-processing systems. Maybe, the reason is that people value decision support systems that respect people's autonomy (Burton et al., 2019). Alternatively, support-after-processing systems leave people in charge of information processing activities, making tasks more demanding and allowing them to show their expertise (Highhouse, 2008). Another possible alternative is that people like decision support systems

that confirm their own, already existing beliefs (Fitzsimons & Lehmann, 2004).

The latter interpretation points towards the task characteristic of feedback as being especially important when considering the current findings. Specifically, while participants did not receive explicit feedback on their task performance, participants who were supported by the support-after-processing system seem to have interpreted the ranking as evidence regarding their performance (Endsley, 2017). This might have led to higher satisfaction with their decisions and also contributed to an increase in self-efficacy through feedback (Renn & Fedor, 2001).

Note that participants in the support-before-processing condition could have also used the support to assess their own decisions and as an indicator of feedback. However, they

Table 6 Summary of hypotheses, contrasts, and results of the contrasts

	NS	SB	SA	df	t	p	d
Hypothesis 1 (performance)	− 2	1	1	See text and Table 7			
Hypothesis 2 (decision time)	1	− 2	1	119	0.53	.30	0.11 [− 0.28, 0.50]
Hypothesis 3a (enjoyment)	1	− 2	1	119	1.01	.15	0.20 [− 0.18, 0.59]
Monotony	− 1	2	− 1	119	0.82	.22	0.17 [− 0.22, 0.56]
Hypothesis 3b (enjoyment)	− 1	–	1	84	1.80	.04	0.40 [− 0 to 03, 0.82]
Monotony	1	–	− 1	84	1.19	.12	0.25 [− 0.18, 0.67]
Hypothesis 4 (satisfaction)	− 1	− 1	2	119	2.47	.01	0.46 [0.09, 0.84]
Hypothesis 5 (responsibility)	1	− 2	1	119	0.67	.25	0.13 [− 0.26, 0.52]
Hypothesis 6a (specific SE)	Interaction contrast: increase from time 1 to time 5 larger for the support-after-processing group than for other groups			119	2.53	.01	0.49 [0.11, 0.87]
Hypothesis 6a (general SE)				119	2.74	.01	0.51 [0.13, 0.89]
Hypothesis 6b (specific SE)	Dependent t test: increase from time 1 to time 5 only for the support-after-processing group and the no-support group		NS	41	0.00	.99	0.00 [− 0.43, 0.43]
			SB	35	2.48	.02	0.41 [− 0.05, 0.88]
			SA	43	4.96	<.001	0.75 [0.31, 1.19]
Hypothesis 6b (general SE)			NS	41	5.85	<.001	0.89 [0.45, 1.35]
			SB	35	3.65	<.001	0.59 [0.12, 1.07]
			SA	43	7.79	<.001	1.16 [0.66, 1.66]

The numbers in the columns NS, SB, and SA reflect the contrast codes for the hypotheses. A coding of 2 or − 2 indicates that this group was compared to the mean of the groups coded with − 1 or 1 (because contrast weights have up to sum up to zero). The positive numbers for the codes indicate that the mean values for the respective groups for the respective dependent variable were expected to be higher. A dash (–) indicates that this group was ignored for the respective comparison

NS no-support group, SB support-before-processing group, SA support-after-processing group, SE self-efficacy

Table 7 Relative frequencies of participants' decisions for the applicants in the given task rounds

	T1			T2			T3			T4			T5			T1–T3 combined		
	C	B	A	C	B	A	C	B	A	C	B	A	C	B	A	C	B	A
No-support	0.00	0.24	0.76	0.00	0.24	0.76	0.00	0.17	0.83	0.00	0.00	1.00	0.00	0.00	1.00	0.00	0.21	0.79
Support-before	0.00	0.08	0.92	0.03	0.25	0.72	0.03	0.11	0.86	0.00	0.06	0.94	0.00	0.03	0.97	0.02	0.15	0.83
Support-after	0.02	0.11	0.87	0.00	0.27	0.73	0.00	0.07	0.93	0.02	0.05	0.93	0.00	0.00	1.00	0.01	0.15	0.84
Support-combined	0.01	0.10	0.89	0.01	0.26	0.73	0.01	0.09	0.90	0.01	0.05	0.94	0.00	0.01	0.99	0.01	0.15	0.84
p Fisher's exact test for Hypothesis 1	.08			.89			.34			–			–			.15		

The numbers in the cell represent relative frequencies for the respective group in a given task round (e.g., in task 1, 24% of participants in the no-support group chose a "B" applicant). Following Hypothesis 1, Fisher's exact test compares the no-support and both conditions receiving support. $n_{\text{no support}} = 42$, $n_{\text{support before}} = 36$, $n_{\text{support after}} = 44$

seem to have interpreted the system's support differently. They might have perceived that there was no real information processing necessary throughout the task because they only had to confirm the systems' recommendation. This seems to be less likely to contribute to decision satisfaction or self-efficacy. This is in line with findings from the area of medical decision-making, where support-before-processing systems have a history of acceptance issues (Arkes et al., 2007; Langlotz & Shortliffe, 1984; Silverman, 1992) because physicians feel reduced to "recipient of the machine's solution" (Silverman, 1992, p. 111). Finding less satisfaction and lower increases in self-efficacy for the support-before-processing system fits into this picture. This finding is also in line with the reluctance to use standardized selection methods (Highhouse, 2008) or with a fear of technological unemployment (Nolan et al., 2016) because of not being able to show own skills when systems already provide possible solutions. In contrast, however, to what could have been expected based on the literature, there were only small and non-significant effects for perceived responsibility, which could be interpreted in a way that all groups might have perceived autonomy during task fulfillment (Chung-Yan, 2010; Hackman & Oldham, 1976). Another possible explanation for this finding is that people were aware that transferring parts of the responsibility to an automated system is not possible (Burton et al., 2019).

Given the differences between the systems, one implication is that implementing different versions of systems differentially affects job motivation, satisfaction, and well-being at work (Burton et al., 2019; Highhouse, 2008). Carefully considering how to implement what kind of systems might therefore be a work design opportunity (Parker & Grote, 2020). Information processing and decision support systems add mechanical information for decision-making that, in previous studies, improved decision quality (Dietvorst et al., 2015; Kuncel et al., 2013). According to our study, certain implementations of such systems have the potential to make jobs more rewarding and provide feelings of individual skill development. We are convinced that how to design and implement automated systems to optimally support human decision-making and well-being at work will be a fruitful avenue for future research. For instance, purposefully considering the timing of when to present outputs of systems to human decision-makers seems to be an important design choice (Sharit, 2003). Based on broader automation literature and tapping into current developments in artificial intelligence (e.g., a focus on eXplainable Artificial Intelligence), we propose further important design choices for future investigation of automated systems for managerial task support to be (a) the automation interface and the presentation of systems' outputs (e.g., ranking, highlighting important applicant information; Endsley, 2017), (b) information and explanations provided together with systems' outputs (e.g., reasons for why certain applicants receive better evaluations; Sokol & Flach, 2020), (c) interactivity of automation design (e.g., providing opportunities to request additional information for outputs; Sheridan & Parasuraman, 2005), and (d) adaptability of automation (e.g., allowing the automated system to take over more tasks vs. restricting its functionalities; Sauer & Chavaillaz, 2018).

Main practical implications

With the implementation of automated decision support systems, organizational systems and employees' jobs change. When organizations want to implement such systems, they should consider potential effects on employees. In personnel selection contexts, one could expect that support-before-processing systems enhance efficiency but might also lead to anchoring effects and hasty decisions (Endsley, 2017). In contrast, support-after-processing systems can enhance work enjoyment and satisfaction with decisions and can improve self-efficacy at work, potentially leading to happier workers. However, such systems do not necessarily enhance efficiency, and their positive effects heavily depend on the reliability of

the system which is likely less than perfect for decision support systems for management tasks.

Limitations and future research

There are at least four limitations that need to be considered and that could guide future research. First, participants were not real hiring managers but people interested in HR management. Actual hiring managers could be more aware of potential consequences of bad hires and have more experience which could affect how they use and react to decision support systems (Highhouse, 2008). However, information processing and decision support systems might especially be considered for novice hiring managers and our results tentatively point towards initial performance benefits of support systems (but see Arnold, Collier, Leech, & Sutton, 2004 who found that decision aids can decrease decision performance for novice decision-makers).

Second, the tasks were not real personnel selection tasks, even if we used actual job descriptions and simulated hiring processes. Therefore, participants did not operate in the complex and stressful world of actual hiring managers. In practice, one could expect different effects as increases in efficiency might be more valuable (and valued) compared to laboratory situations. Specifically, in real job contexts, following recommendations by automated systems might be more tempting as this can drastically decrease time to reach decisions and thus makes more time available for other tasks (Endsley, 2017).

Third, the current study only investigated decision support systems for personnel selection. This is just one part of managers' work (although a large part of hiring managers' work). Therefore, effects on general managers' psychological states at work might be small. However, there are other contexts within organizations where automated systems based on artificial intelligence are on the edge of changing work processes (e.g., for scheduling, turnover prediction; Ötting & Maier, 2018; Sajjadiani et al., 2019). For these contexts, questions similar to the ones we examined in the current paper evolve: how does the implementation of automated systems affect psychological states and performance at work and how to optimally design and implement systems to foster employee well-being? Additionally, we imagine that there are task- and context-specific characteristics that affect how people use and how they react to systems. For instance, future research could manipulate time pressure and/or investigate reactions to decision support systems in multitasking environments (see, for instance, Karpinsky, Chancey, Palmer, and Yamani (2018) and Wiczorek and Manzey (2014) for examples of such studies). Participants could, for instance, receive bonus for fulfilling as many as possible tasks which could lead to more blindly following recommendations by automated systems.

Fourth, participants hitting the performance ceiling after task 3 in our design might undermine the generalizability of the results for task performance. Apparently, participants quickly learned the task which is also reflected by the fact that they completed the tasks more quickly over time and by the finding that task enjoyment decreased, while monotony increased. This learning effect might have reduced variance in task performance to a point that performance benefits from the systems are concealed by lack of variance. Future research should therefore increase task complexity (e.g., by adding further information) in order to investigate the generalizability of our results. Based on a body of previous research, it might be expectable that there are performance benefits of using automated systems (Burton et al., 2019; Kuncel et al., 2013). However, it might be something quite different to show benefits of mechanically combining information from different sources of selection methods versus relying on tools based on artificial intelligence and machine learning that automatically evaluate human behavior. For automated systems for managerial tasks arising from the emerging use of artificial intelligence, research on advantages and disadvantages regarding decision-making, trust in those systems, and an optimal design to use them at work is still missing—the current study is an initial step to fill this gap.

Conclusion

We present a first attempt to investigate how systems from a new era of human-automation interaction may change existing management tasks and affect managerial jobs. Additionally, we reinforce calls for research investigating how to design optimal collaboration between humans and automated systems. By marrying human factors and management literature, psychology can get involved into the debate on artificial intelligence and automation in management. This could expand the focus from effectiveness and efficiency questions towards a broader work design frame and towards psychological states that might be affected by changing work environments through the implementation of management-automation systems.

References

Ambady, N., Krabbenhoft, M. A., & Hogan, D. (2006). The 30-sec sale: Using thin-slice judgments to evaluate sales effectiveness. *Journal of Consumer Psychology, 16*(1), 4–13. https://doi.org/10.1207/s15327663jcp1601_2.

Arkes, H. R., Shaffer, V. A., & Medow, M. A. (2007). Patients derogate physicians who use a computer-assisted diagnostic aid. *Medical Decision Making, 27*(2), 189–202. https://doi.org/10.1177/0272989X06297391.

Arnold, V., Collier, P. A., Leech, S. A., & Sutton, S. G. (2004). Impact of intelligent decision aids on expert and novice decision-makers' judgments. *Accounting and Finance, 44*(1), 1–26. https://doi.org/10.1111/j.1467-629x.2004.00099.x.

Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review, 84*(2), 191–215. https://doi.org/10.1037//0033-295x.84.2.191.

Brehaut, J. C., O'Connor, A. M., Wood, T. J., Hack, T. F., Siminoff, L., Gordon, E., & Feldman-Stewart, D. (2003). Validation of a decision regret scale. *Medical Decision Making, 23*(4), 281–292. https://doi.org/10.1177/0272989X03256005.

Brennan, T., Dieterich, W., & Ehret, B. (2009). Evaluating the predictive validity of the COMPAS risk and needs assessment system. *Criminal Justice and Behavior, 36*(1), 21–40. https://doi.org/10.1177/0093854808326545.

Bright, T. J., Wong, A., Dhurjati, R., Bristow, E., Bastian, L., Coeytaux, R. R., Samsa, G., Hasselblad, V., Williams, J. W., Musty, M. D., Wing, L., Kendrick, A. S., Sanders, G. D., & Lobach, D. (2012). Effect of clinical decision support systems: A systematic review. *Annals of Internal Medicine, 157*(1), 29–43. https://doi.org/10.7326/0003-4819-157-1-201207030-00450.

Burton, J. W., Stein, M., & Jensen, T. B. (2019). A systematic review of algorithm aversion in augmented decision making. *Journal of Behavioral Decision Making*. Advance online publication, *33*, 220–239. https://doi.org/10.1002/bdm.2155.

Campion, M. C., Campion, M. A., Campion, E. D., & Reider, M. H. (2016). Initial investigation into computer scoring of candidate essays for personnel selection. *Journal of Applied Psychology, 101*(7), 958–975. https://doi.org/10.1037/apl0000108.

Chung-Yan, G. A. (2010). The nonlinear effects of job complexity and autonomy on job satisfaction, turnover, and psychological well-being. *Journal of Occupational Health Psychology, 15*(3), 237–251. https://doi.org/10.1037/a0019823.

Crosslands, M. D., Wynne, B. E., & Perkins, W. C. (1995). Spatial decision support systems: An overview of technology and a test of efficacy. *Decision Support Systems, 14*(3), 219–235. https://doi.org/10.1016/0167-9236(94)00018-N.

de Visser, E. J., Pak, R., & Shaw, T. H. (2018). From 'automation' to 'autonomy': The importance of trust repair in human–machine interaction. *Ergonomics, 61*, 1409–1427. https://doi.org/10.1080/00140139.2018.1457725.

Derous, E., Buijsrogge, A., Roulin, N., & Duyck, W. (2016). Why your stigma isn't hired: A dual-process framework of interview bias. *Human Resource Management Review, 26*(2), 90–111. https://doi.org/10.1016/j.hrmr.2015.09.006.

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. https://doi.org/10.1037/xge0000033.

Endsley, M. R. (2017). From here to autonomy: Lessons learned from human–automation research. *Human Factors, 59*(1), 5–27. https://doi.org/10.1177/0018720816681350.

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

Fitzsimons, G. J., & Lehmann, D. R. (2004). Reactance to recommendations: When unsolicited advice yields contrary responses. *Marketing Science, 23*(1), 82–94. https://doi.org/10.1287/mksc.1030.0033.

Gonzalez, M. F., Capman, J. F., Oswald, F. L., Theys, 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), 5. https://doi.org/10.25035/pad.2019.03.005.

Guerlain, S. A., Smith, P. J., Obradovich, J. H., Rudmann, S., Strohm, P., Smith, J. W., Svirbely, J., & Sachs, L. (1999). Interactive critiquing as a form of decision support: An empirical evaluation. *Human Factors, 41*(1), 72–89. https://doi.org/10.1518/001872099779577363.

Hackman, J. R., & Oldham, G. R. (1976). Motivation through the design of work: Test of a theory. *Organizational Behavior and Human Performance, 16*(2), 250–279. https://doi.org/10.1016/0030-5073(76)90016-7.

Hiemstra, A. M. F., Oostrom, J. K., Derous, E., Serlie, A. W., & Born, M. P. (2019). Applicant perceptions of initial job candidate screening with asynchronous job interviews: Does personality matter? *Journal of Personnel Psychology, 18*(3), 138–147. https://doi.org/10.1027/1866-5888/a000230.

Highhouse, S. (2008). Stubborn reliance on intuition and subjectivity in employee selection. *Industrial and Organizational Psychology, 1*(3), 333–342. https://doi.org/10.1111/j.1754-9434.2008.00058.x.

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.

Houston, D. A., Sherman, S. J., & Baker, S. M. (1991). Feature matching, unique features, and the dynamics of the choice process: Predecision conflict and postdecision satisfaction. *Journal of Experimental Social Psychology, 27*(5), 411–430. https://doi.org/10.1016/0022-1031(91)90001-M.

Humphrey, S. E., Nahrgang, J. D., & Morgeson, F. P. (2007). Integrating motivational, social, and contextual work design features: A meta-analytic summary and theoretical extension of the work design literature. *Journal of Applied Psychology, 92*(5), 1332–1356. https://doi.org/10.1037/0021-9010.92.5.1332.

Jerusalem, M., & Schwarzer, R. (1992). Self-efficacy as a resource factor in stress appraisal processes. In R. Schwarzer (Ed.), *Self-efficacy: Thought control of action* (pp. 195–216). Routledge.

Judge, T. A., Jackson, C. L., Shaw, J. C., Scott, B. A., & Rich, B. L. (2007). Self-efficacy and work-related performance: The integral role of individual differences. *Journal of Applied Psychology, 92*(1), 107–127. https://doi.org/10.1037/0021-9010.92.1.107.

Karpinsky, N. D., Chancey, E. T., Palmer, D. B., & Yamani, Y. (2018). Automation trust and attention allocation in multitasking workspace. *Applied Ergonomics, 70*, 194–201. https://doi.org/10.1016/j.apergo.2018.03.008.

König, C. J., Klehe, U.-C., Berchtold, M., & Kleinmann, M. (2010). Reasons for being selective when choosing personnel selection procedures. *International Journal of Selection and Assessment, 18*(1), 17–27. https://doi.org/10.1111/j.1468-2389.2010.00485.x.

Kuncel, N. R. (2018). Judgment and decision making in staffing research and practice. In D. Ones, N. Anderson, C. Viswesvaran, & H. Sinangil (Eds.), *The Sage handbook of industrial, work and organizational psychology: Personnel psychology and employee performance* (pp. 474–487). Sage. https://doi.org/10.4135/9781473914940.n17.

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://doi.org/10.1037/a0034156.

Langer, M., König, C. J., & Hemsing, V. (2020). Is anybody listening? The impact of automatically evaluated job interviews on impression

management and applicant reactions. *Journal of Managerial Psychology, 35*(4), 271-284. https://doi.org/10.1108/JMP-03-2019-0156.

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

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

Langlotz, C. P., Allen, B., Erickson, B. J., Kalpathy-Cramer, J., Bigelow, K., Cook, T. S., Flanders, A. E., Lungren, M. P., Mendelson, D. S., Rudie, J. D., Wang, G., & Kandarpa, K. (2019). A roadmap for foundational research on artificial intelligence in medical imaging. *Radiology, 291*(3), 781–791. https://doi.org/10.1148/radiol.2019190613.

Langlotz, C. P., & Shortliffe, E. H. (1984). Adapting a consultation system to critique user plans. *International Journal of Man-Machine Studies, 19*(5), 479–496. https://doi.org/10.1016/s0020-7373(83)80067-4.

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.

Lent, R. W., Brown, S. D., & Larkin, K. C. (1987). Comparison of three theoretically derived variables in predicting career and academic behavior: Self-efficacy, interest congruence, and consequence thinking. *Journal of Counseling Psychology, 34*(3), 293–298. https://doi.org/10.1037/0022-0167.34.3.293.

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.

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

Lowe, D. J., Reckers, P. M. J., & Whitecotton, S. M. (2002). The effects of decision-aid use and reliability on jurors' evaluations of auditor liability. *Accounting Review, 77*(1), 185–202. https://doi.org/10.2308/accr.2002.77.1.185.

Luszczynska, A., Scholz, U., & Schwarzer, R. (2005). The general self-efficacy scale: Multicultural validation studies. *Journal of Psychology, 139*(5), 439–457. https://doi.org/10.3200/JRLP.139.5.439-457.

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.

McKnight, D. H., Carter, M., Thatcher, J. B., & Clay, P. F. (2011). Trust in a specific technology: An investigation of its components and measures. *ACM Transactions on Management Information Systems, 2*(2), 1–25. https://doi.org/10.1145/1985347.1985353.

Meehl, P. E. (1954). *Clinical versus statistical prediction: A theoretical analysis and a review of the evidence.* University of Minnesota Press.

Morgeson, F. P., Garzsa, A. S., & Campion, M. A. (2012). Work design. In E. B. Weiner, N. W. Schmitt, & S. Highhouse (Eds.), *Handbook of psychology* (pp. 525–559). Wiley. https://doi.org/10.1002/9781118133880.hop212020.

Morgeson, F. P., & Humphrey, S. E. (2006). The Work Design Questionnaire (WDQ): Developing and validating a comprehensive measure for assessing job design and the nature of work. *Journal of Applied Psychology, 91*(6), 1321–1339. https://doi.org/10.1037/0021-9010.91.6.1321.

Muir, B. M. (1994). Trust in automation: Part I. Theoretical issues in the study of trust and human intervention in automated systems.

Ergonomics, 37(11), 1905–1922. https://doi.org/10.1080/00140139408964957.

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

Nolan, K., Carter, N., & Dalal, D. (2016). Threat of technological unemployment: Are hiring managers discounted for using standardized employee selection practices? *Personnel Assessment and Decisions, 2*(1), 4. https://doi.org/10.25035/pad.2016.004.

Oldham, G. R., & Fried, Y. (2016). Job design research and theory: Past, present and future. *Organizational Behavior and Human Decision Processes, 136*, 20–35. https://doi.org/10.1016/j.obhdp.2016.05.002.

Onnasch, L., Wickens, C. D., Li, H., & Manzey, D. (2014). Human performance consequences of stages and levels of automation: An integrated meta-analysis. *Human Factors, 56*(3), 476–488. https://doi.org/10.1177/0018720813501549.

Ötting, S. K., & Maier, G. W. (2018). The importance of procedural justice in human-machine-interactions: Intelligent systems as new decision agents in organizations. *Computers in Human Behavior, 89*, 27–39. https://doi.org/10.1016/j.chb.2018.07.022.

Parasuraman, R., & Manzey, D. H. (2010). Complacency and bias in human use of automation: An attentional integration. *Human Factors, 52*(3), 381–410. https://doi.org/10.1177/0018720810376055.

Parasuraman, R., Sheridan, T. B., & Wickens, C. D. (2000). A model for types and levels of human interaction with automation. *IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans, 30*(3), 286–297. https://doi.org/10.1109/3468.844354.

Parker, S. K., & Grote, G. (2020). Automation, algorithms, and beyond: Why work design matters more than ever in a digital world. *Applied Psychology.* Advance online publication. https://doi.org/10.1111/apps.12241.

Plath, H.-E., & Richter, P. (1984). *Ermüdung-Monotonie-Sättigung-Stress: Verfahren zur skalierten Erfassung erlebter Beanspruchungsfolgen* [Fatigue-monotony-saturation-stress: A measure to capture experienced stress results]. Hogrefe.

Raghavan, M., Barocas, S., Kleinberg, J., & Levy, K. (2020). Mitigating bias in algorithmic hiring: Evaluating claims and practices. In *FAT* '20: Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency* (pp. 469–481). https://doi.org/10.1145/3351095.3372828.

Renn, R. W., & Fedor, D. B. (2001). Development and field test of a feedback seeking, self-efficacy, and goal setting model of work performance. *Journal of Management, 27*(5), 563–583. https://doi.org/10.1177/014920630102700504.

Robertson, I. T., & Smith, M. (2001). Personnel selection. *Journal of Occupational and Organizational Psychology, 74*(4), 441–472. https://doi.org/10.1348/096317901167479.

Sainfort, F., & Booske, B. C. (2000). Measuring post-decision satisfaction. *Medical Decision Making, 20*(1), 51–61. https://doi.org/10.1177/0272989X0002000107.

Sajjadiani, S., Sojourner, A. J., Kammeyer-Mueller, J. D., & Mykerezi, E. (2019). Using machine learning to translate applicant work history into predictors of performance and turnover. *Journal of Applied Psychology, 104*(10), 1207–1225. https://doi.org/10.1037/apl0000405.

Sansone, C. (1989). Competence feedback, task feedback, and intrinsic interest: An examination of process and context. *Journal of Experimental Social Psychology, 25*(4), 343–361. https://doi.org/10.1016/0022-1031(89)90027-9.

Sauer, J., & Chavaillaz, A. (2018). How operators make use of wide-choice adaptable automation: Observations from a series of

experimental studies. *Theoretical Issues in Ergonomics Science, 19*(2), 135–155. https://doi.org/10.1080/1463922X.2017.1297866.

Schwarzer, R., & Jerusalem, M. (1995). Generalized self-efficacy scale. In J. Weinman, S. Wright, & M. Johnston (Eds.), *Measures in health psychology: A user's portfolio causal and control beliefs* (pp. 35–37). NFER-NELSON.

Schyns, B., & von Collani, G. (2014). Berufliche Selbstwirksamkeitserwartung [Occupational self-efficacy expectations]. Zusammenstellung sozialwissenschaftlicher Items und Skalen (ZIS). https://doi.org/10.6102/zis16.

Sharit, J. (2003). Perspectives on computer aiding in cognitive work domains: Toward predictions of effectiveness and use. *Ergonomics, 46*(1–3), 126–140. https://doi.org/10.1080/00140130303533.

Sheridan, T. B. (2000). Function allocation: Algorithm, alchemy or apostasy? *International Journal of Human-Computer Studies, 52*(2), 203–216. https://doi.org/10.1006/ijhc.1999.0285.

Sheridan, T. B., & Parasuraman, R. (2005). Human automation interaction. *Reviews of Human Factors, 1*(1), 89–129. https://doi.org/10.1518/155723405783703082.

Silverman, B. C. (1992). Survey of expert critiquing systems: Practical and theoretical frontiers. *Communications of the ACM, 35*(4), 106–127. https://doi.org/10.1145/129852.129861.

Smith, M. J., & Carayon, P. (1995). New technology, automation, and work organization: Stress problems and improved technology implementation strategies. *International Journal of Human Factors in Manufacturing, 1*, 99–116. https://doi.org/10.1002/hfm.4530050107.

Smith, R. P. (1981). Boredom: A review. *Human Factors, 23*(3), 329–340. https://doi.org/10.1177/001872088102300308.

Sokol, K., & Flach, P. (2020). Explainability fact sheets: A framework for systematic assessment of explainable approaches. In *FAT* '20: Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency* (pp. 56–67). https://doi.org/10.1145/3351095.3372870.

Stajkovic, A. D., & Luthans, F. (1998). Self-efficacy and work-related performance: A meta-analysis. *Psychological Bulletin, 124*(2), 240–261. https://doi.org/10.1037/0033-2909.124.2.240.

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

Taber, T. D., & Alliger, G. M. (1995). A task-level assessment of job satisfaction. *Journal of Organizational Behavior, 16*(2), 101–121. https://doi.org/10.1002/job.4030160202.

Trippel, C. (2012). *Selbstwirksamkeit von selbststeuernden Teams und ihren unmittelbaren Vorgesetzten: Zusammenhänge mit Konflikten, Zusammenarbeit, Führung und Leistung* [Self-efficacy of self-managing teams and their direct leaders: Relationships with conflicts, teamwork, leadership, and performance] (Doctoral thesis). Universität Mannheim. https://madoc.bib.uni-mannheim.de/29975/1/Trippel_Dissertation.pdf

van Dongen, K., & van Maanen, P.-P. (2013). A framework for explaining reliance on decision aids. *International Journal of Human-Computer Studies, 71*(4), 410–424. https://doi.org/10.1016/j.ijhcs.2012.10.018.

Warr, P. (1994). A conceptual framework for the study of work and mental health. *Work & Stress, 8*(2), 84–97. https://doi.org/10.1080/02678379408259982.

Wegman, L. A., Hoffman, B. J., Carter, N. T., Twenge, J. M., & Guenole, N. (2018). Placing job characteristics in context: Cross-temporal meta-analysis of changes in job characteristics since 1975. *Journal of Management, 44*(1), 352–386. https://doi.org/10.1177/0149206316654545.

Wesche, J. S., & Sonderegger, A. (2019). When computers take the lead: The automation of leadership. *Computers in Human Behavior, 101*, 197–209. https://doi.org/10.1016/j.chb.2019.07.027.

Wiczorek, R., & Manzey, D. (2014). Supporting attention allocation in multitask environments: Effects of likelihood alarm systems on trust, behavior, and performance. *Human Factors, 56*(7), 1209–1221. https://doi.org/10.1177/0018720814528534.