

Give Me a Choice: The Consequences of Restricting Choices Through AI-Support for Perceived Autonomy, Motivational Variables, and Decision Performance

CEDRIC FAAS, Saarland Informatics Campus, Saarland University, Germany

RICHARD BERGS, Department of Psychology, University of Freiburg, Germany

SARAH STERZ, Saarland Informatics Campus, Saarland University, Germany

MARKUS LANGER*, Department of Psychology, University of Freiburg, Germany

ANNA MARIA FEIT*, Saarland Informatics Campus, Saarland University, Germany

Design optimizations in human-AI collaboration often focus on cognitive aspects like attention and task load. Drawing on work design literature, we propose that effective human-AI collaboration requires broader consideration of human needs (e.g., autonomy) that affect motivational variables (e.g., meaningfulness). In a simulated drone oversight experiment, participants ($N=274$, between-subject) faced 10 critical decision-making scenarios with varying levels of choice restrictions with an AI recommending only 1, 2, 4 or all 6 possible actions. Restricting participants to one selectable action improved task performance (with a perfect AI) but significantly reduced perceived autonomy and work meaningfulness, and these effects intensified over time. In conditions with multiple action choices, participants with higher perceived autonomy performed better. The findings underscore the importance of considering motivational factors to design successful long-term human-AI collaboration at work.

CCS Concepts: • **Human-centered computing** → **Empirical studies in HCI**; **User studies**;

Additional Key Words and Phrases: autonomy, psychological needs, motivation, human oversight, AI decision support

1 Introduction

AI-based systems can support users in complex or time-sensitive scenarios by providing for example decision recommendations [5, 9, 29, 70, 95]. Typically, the goal of such systems is to improve overall task performance (e.g., decision accuracy, decision time). Thus, most research has focused on optimizing the accuracy of AI methods [53, 54, 56, 63, 72] or the joint human-AI performance, by considering cognitive factors such as users' attention [17, 62], cognitive load [24, 36, 39], or their understanding of AI recommendations [4, 5, 45].

However, there is a history of research in work psychology showing that effectiveness in work tasks, including decision-making, depends not only on human cognition but also on human motivation [13, 34, 68], which are closely linked to considering psychological needs (e.g., a need for autonomy, a need for competence; [73]). Regarding the design of AI-based systems to support humans at work, we see only few works so far that consider the influence of motivational aspects on worker satisfaction and performance [6, 8, 18]. Research on human-AI interactions still mostly emphasizes cognitive and performance-related aspects. Problematically, when focusing predominantly on optimizing for cognitive aspects, this can even undermine motivational aspects of work. For example, in an attempt to reduce users' cognitive load, designers may consider implementing AI-based decision support that reduces the choices that users have for actions. Simultaneously, such a restriction of choices can reduce users' perceived autonomy potentially

*Both authors supervised this research equally.

Authors' Contact Information: Cedric Faas, s8cefaas@uni-saarland.de, Saarland Informatics Campus, Saarland University, Saarbrücken, Germany; Richard Bergs, richard.bergs@psychologie.uni-freiburg.de, Department of Psychology, University of Freiburg, Freiburg im Breisgau, Germany; Sarah Sterz, sterz@depend.uni-saarland.de, Saarland Informatics Campus, Saarland University, Saarbrücken, Germany; Markus Langer, markus.langer@psychologie.uni-freiburg.de, Department of Psychology, University of Freiburg, Freiburg im Breisgau, Germany; Anna Maria Feit, feit@cs.uni-saarland.de, Saarland Informatics Campus, Saarland University, Saarbrücken, Germany.

undermining people’s perceived meaningfulness of their task with meaningfulness being one central variable related to long-term job motivation and satisfaction [68]. Ironically, this can then also negatively impact long-term human-AI joint performance given the importance of motivation and satisfaction for job performance [13]. Thus, we argue that human-AI interaction also needs to be optimized considering motivational aspects.

Motivated by theories from the work design literature (e.g., [34, 57, 68]), this paper aims to establish a relationship between psychological needs, motivational variables and users’ performance when interacting with AI systems at work. Specifically, our goal is to explore how the restriction of choices through AI systems affects the psychological need for autonomy, and how this relates to users’ motivation and task performance in AI-supported decision-making tasks. We present results from an experimental study that tested the consequences of reducing people’s choices for actions in a simulated drone oversight task. Participants were in the role of a human overseer of an autonomous drone that faced different critical situations that required the human overseer to decide for an action on how the drone can safely resolve the situation. The four between-participant experimental conditions were that the human overseer received decision support by an AI-based system that offered six actions to resolve the situation, or that of those *Six* actions only *Four* vs. *Two* vs. *One* were selectable while the other options were grayed out (i.e. not recommended by the AI). Following theories in work design, we hypothesized that restricting the number of *Selectable Actions* would affect participants’ perceived autonomy. We expected that this will impact downstream motivational aspects of work such as perceived meaningfulness, with potential effects on human task performance in the oversight task.

Our findings revealed that having only *One Selectable Action* decreased participants perceived autonomy and perceived meaningfulness of the task compared to being able to choose between *Two* or more actions (while increasing the decision accuracy, since the remaining action was always correct). This effect increased over time. In the conditions that had more than only *One Selectable Action*, participants who felt more autonomous were also the ones who showed a better decision accuracy.

We see three main contributions of this paper. First, we highlight the importance of considering motivational aspects in designing beneficial and effective human-AI interactions and highlight the value of the work design literature for this endeavor. Second, we show that whereas restricting participants’ choices during the oversight task increased their task performance (an effect that strongly depends on the accuracy of the AI-based decision support), this can also reduce their perceived autonomy and perceived meaningfulness of their work. These effects on such motivational aspects of work grew stronger over time showing the need to examine human-AI interactions in the longer term. Third, our findings emphasize the importance of evaluating the effects of motivational aspects on joint human-AI performance at work given that participants who felt more autonomous also performed better than those who felt less autonomous.

2 Related Work

2.1 AI-supported decision-making

Research on AI-supported decision-making covers many domains [44]. For example, AI-based systems can support decision-making in domains such as education [71], finance [1, 3, 14, 95], healthcare [3, 11, 12, 31, 41], hiring [45, 49, 55] or law [30, 42, 52, 91].

One key goal of AI-supported decision-making is to foster effective decisions. The effectiveness of human-AI collaborations can be measured, for instance, by the joint decision accuracy of humans and AI systems [53, 54, 56, 63, 72, 75, 83]. Other measures of effectiveness include the time required for decision-making with AI-based systems,

especially in time-critical decision contexts [53, 56, 71, 83]. In theory, effective human-AI collaboration is possible, and studies show that AI-based systems can improve decision-making across various contexts [44].

However, recent studies indicate that human-AI collaborations can also lead to suboptimal outcomes, as the joint performance of humans and systems may often be worse than fully automating tasks with an AI-based system [5, 9, 29, 70, 95]. For example, users of AI-based systems may find it difficult to discern accurate from inaccurate system outputs [27, 30, 94], and, as a consequence, they may override actually accurate system outputs and follow inaccurate ones instead [28].

Thus, research aims to make human-AI collaborations more effective. A large share of attempts to optimize human-AI decision-making to date has been focused on cognitive aspects relating to human decision-making performance. For example, the area of XAI tries to achieve better joint performance by enhancing people’s understanding of AI-based decision processes and outputs [38, 48, 50, 72, 92]. This should help users decide when to follow, reject, or adapt system outputs [10, 75]. Other research aims to improve joint decision-making by designing systems that can communicate uncertainty in their outputs, ultimately targeting a similar goal as explanations: providing users with better insights into when to follow or reject system outputs [5]. In an attempt to encourage more thoughtful decision-making, further research suggests using cognitive forcing methods, such as altering the timing of when users receive AI-based outputs (e.g., only after users have made an initial decision [10, 46]), or introducing a time lag until people receive the AI-based output [8]. Other research focuses on optimizing users’ cognitive load, for example, by automatically adapting interfaces to their current workload [51] or reducing the perceptual load of users by changing the presentation of information-dense interfaces [24].

While optimizing cognitive aspects in human-AI decision-making is crucial, work psychology literature reviewed in the next section emphasizes the role of psychological needs and motivational factors for worker effectiveness and satisfaction. In line with emerging research highlighting the importance of considering motivational aspects for optimizing human-AI interaction [6, 80], in this paper we want to explore in how far this generalizes to work supported by AI-based systems.

2.2 Motivational Aspects of human-AI Collaboration at Work

Work psychology literature emphasizes that effectiveness in work tasks, such as decision-making, depends not only on human cognition but also on human motivation [13, 34, 57, 68]. In human-AI interaction, these motivational aspects have only recently begun to receive attention [6, 80] with researchers such as Steyvers and Kumar [80] calling for studies to explore the impact of motivational factors, particularly on the long-term use of AI tools. Therefore, in the following, we focus on work psychology literature to understand the relationship between psychological needs, motivation, and performance, which has a long history of research in the work context (e.g., [34]).

The importance of motivational aspects of work can be understood by considering the core functions of human motivation: Motivation gives direction, that is which behavior one chooses to perform, it affects the degree of effort one invests in a behavior, and it impacts the persistence and duration of that effort [13, 90]. All else being equal, a motivated worker should therefore perform better for a longer period of time than a less motivated worker. Furthermore, research typically expects and finds a strong relation between motivation and satisfaction [40, 57, 68], meaning that motivated workers are also expected to be more satisfied with their work for a longer period of time.

There are various streams of research that help us understand which factors in human-AI collaboration might contribute to motivational aspects of work. These streams share a common focus on describing *psychological needs* and *psychological processes* that contribute to motivating work environments.

Self-Determination Theory is one of the broadest and most comprehensive theories describing three basic psychological needs that contribute to human motivation: a need for autonomy, competence, and relatedness ([73]). The need for autonomy refers to having choices and being the initiator of one's own actions; the need for competence pertains to engaging in optimally challenging tasks and effectively achieving desired outcomes; the need for relatedness refers to feeling a sense of belonging and building relationships with others. In the workplace, autonomy is positively associated with motivation and well-being [40, 88]. Autonomy can also positively affect work performance and effort [15, 87]. Beyond the work context, research shows positive effects of autonomy on positive affect and well-being [74, 78, 84].

Most of what is proposed by the Self-Determination Theory is consistent with the key propositions in the work design literature. While there are many theories and models to draw from (see, e.g., [33, 34, 57, 58]), we decided to build on the SMART model by Parker & Knight [68] to further understand how fulfilling psychological needs can contribute to effective human-AI collaborations. The *SMART model* refers to five different characteristics of work that influence work satisfaction through different psychological processes: Stimulating work characteristics, Mastery of work characteristics, Autonomous work characteristics, Relational work characteristics, and Tolerable work characteristics. In this paper we particularly focus on the effects of autonomous work characteristics that are supposed to affect users primarily through motivational aspects. Specifically, Parker & Knight argue that having choices and control with respect to the decisions one makes at work and with respect to how to do their work will lead to a higher perceived autonomy and, consequently, a higher perceived work meaningfulness [68]. In the long term, this is expected to contribute to a higher motivation at work and to a higher job satisfaction. These propositions are supported by meta-analyses of work psychology studies showing that autonomy relates to variables such as meaningfulness, motivation, and job satisfaction [40].

Interdisciplinary research on the effects of implementing AI at work implies that AI-based systems may strongly affect people's autonomy [64, 67]. For example, research on the algorithmic management of Uber drivers and platform workers shows that AI-based tools managing workers can undermine workers' perceived feelings of being in control of how they can perform their work [47, 60]. Ulfert et al. found that a system's level of automation influenced users' intention to use the system, but did not find evidence that high-autonomy decision support systems increase stress [86]. Findings from a study by Passalacqua et al. indicate that fully automating the selection of optimal decisions, compared to involving users in the decision selection, negatively impacted their perceived autonomy and motivation during the training for an assembly line task [69]. In a case study by Strich et al. [82], loan officers, whose primary task was to input information into an AI-based system and communicate the AI-based decision to loan applicants, were found to start manipulating the input data to be able to affect the AI-based outputs and thus regain their autonomy over the AI-controlled decision processes. De Vreede et al. [18] conducted a study where participants with always available AI-support felt less autonomous but performed better than those who could choose whether the AI-support was revealed for each decision. We suspect that the effect on decision accuracy in the experiment by De Vreede et al. was due to the experimental conditions rather than the participants' perceived autonomy. Specifically, the experimental conditions were designed in a way that could make deciding correctly easier if the AI-support was always available. In the condition where participants had to turn on the AI-support, decisions might have been worse because the (mostly accurate) support was used less. This highlights the importance of analyzing the effects of perceived autonomy on decision accuracy while properly accounting for potential effects of the accuracy of available AI-support.

3 Method

Motivated by the work design literature, the goal of our research was to explore which role psychological needs and motivational processes play in the design and use of AI-decision support systems at work. Given the research in this field discussed above, we were particularly interested in perceived autonomy as one of three major psychological needs, and how a restriction in users' perceived autonomy might impact motivational aspects such as perceived meaningfulness and in conclusion users' task performance.

Therefore, we devised an online experiment where participants were tasked to monitor an autonomously flying delivery drone and its sensor values and quickly decide on suitable actions when the drone encountered critical situations (e.g., low battery or strong winds).

We chose this task because it represented a realistic scenario for decision-making under time constraints that participants would be able to relate to. At the same time it allowed us to measure the task performance as a result of the drone crashing or being saved. To integrate AI recommendations and simplify the task for participants inexperienced in drone flying, the monitoring interface provided users with six actions to choose from for how to resolve the critical situation (e.g. landing the drone or flying higher) rather than giving participants full control over the drone's flight. This enabled us to easily integrate different levels of (simulated) AI support. In three of the four experimental conditions, the AI-based system provided additional decision support by graying out certain actions that the AI would not recommend. Given that the grayed-out actions were also not selectable anymore, this restricted participants in their choice and thus their autonomy.

In the following, we describe our research questions, study design, and data collection. Our study was approved by the department's ethical review board and preregistered via the [Open Science Foundation](#), where we also published the collected data.

3.1 Research Questions

Based on our theoretical reasoning informed by the Self-Determination Theory and the work design literature, we expect that the degree of restriction people experience through AI-support in a decision-making task will affect users in multiple ways. First, we expect that restricting users' choices will negatively affect users' perceived autonomy and will detrimentally affect downstream motivational effects on users' perceived meaningfulness, task motivation, and task satisfaction. These motivational effects may also negatively impact users' overall task performance.

However, decreasing the number of *Selectable Actions* in a decision-making task can also reduce users' task load [7, 16]. Additionally, in case of a reliable and highly accurate AI-based systems, having few *Selectable Actions* can also directly contribute to increased joint human-AI decision accuracy given that actions that would lead to unfavorable outcomes will not be recommended by the system (see also [18]). Therefore, reducing the number of *Selectable Actions* might positively affect users' decision accuracy and reduce the task load they experience.

We thus face contradicting possible effects of restricting users' choices in the decision making task: restricting choices may reduce users' perceived autonomy which can detrimentally affect motivational aspects, thus also potentially decreasing task performance; but restricting choices may also reduce users' task load or may have a direct positive influence on task performance by reducing the number of ineffective available choices. Given these contradicting effects that could play out differently in our experimental conditions, we refrained from proposing directed hypotheses for our experimental conditions, and instead ask the following research questions:

Does the difference in the number of *Selectable Actions* affect ...

- RQ1a. ...users' perceived autonomy?
- RQ1b. ...users' perceived meaningfulness?
- RQ1c. ...users' task motivation?
- RQ1d. ...users' task load?
- RQ1e. ...the accuracy of decisions made in the drone handover task?
- RQ1f. ...users' decision time?
- RQ1g. ...users' task satisfaction?

Additionally, we explore whether perceived autonomy, meaningfulness, task motivation, and task satisfaction are related to ...

RQ2a. ...the accuracy of decisions made in the drone handover task.

RQ2b. ...users' decision time.

3.2 Participants

We conducted an a-priori calculation of the required sample size with G*Power [23]. We determined a required sample size of $N = 271$, to achieve a power of $1 - \text{BETA} = .95$ given an ALPHA level of .05 and expecting a medium effect size of $f = .25$ in a between-participants ANOVA with four conditions.¹ We decided to conduct the study as an online experiment on Prolific, where we restricted the sampling to English-speaking participants, participants with a prior approval rate greater than 95%, and who joined Prolific more than a year ago. We collected data from 280 participants (151 female, 129 male, age: 18-71, $M = 33.25$, $SD = 9.81$). The participants were compensated with £6.32 for participating in the study (median completion time: 40 minutes, this amounts to an hourly wage of £9.48).

In line with our preregistered exclusion criteria, we excluded four participants who stated that their data should not be used, and two participants who experienced technical issues. There were no participants who showed clear signs of inattentive responding (e.g., all participants took longer for their participation than our preregistered exclusion criterion of 15 minutes).

3.2.1 Study design. The experiment was a between-participants study with *Selectable Actions* as the experimental factor. The different levels of *Selectable Actions* are *One*, *Two*, *Four*, and *Six Selectable Actions*. For each handover scenario, the participants saw six possible actions, but depending on the experimental condition some of them were grayed out and not selectable, as shown in Figure 1. The action that successfully resolved the situation was always available. Accordingly, in condition *One*, the only *Selectable Action* would always resolve the critical situation successfully. Actions selectable in conditions *Four* and *Two* were selected based on outcomes from the study performed by Gundappa et al. [32], where the other *Selectable Actions* were the ones most often chosen by participants.

3.3 Task and Procedure

Participants watched ten 40s-videos of simulated drones flying autonomously. Their task was to monitor the camera output and displayed sensor values to make an informed choice at the end of each video, when participants had to take control of the drone during a critical situation where they then had to choose a suitable action that would avoid a crash.

¹Note that consistent with our preregistration, we finally decided to use linear mixed models for most of the analyses given that they can efficiently analyze and communicate the direction of main effects and interaction effects of independent variables. Given the similarity between ANOVA analyses and regression analyses such as linear mixed models, these calculations should also be suited to approximate the required number of participants for linear mixed models.

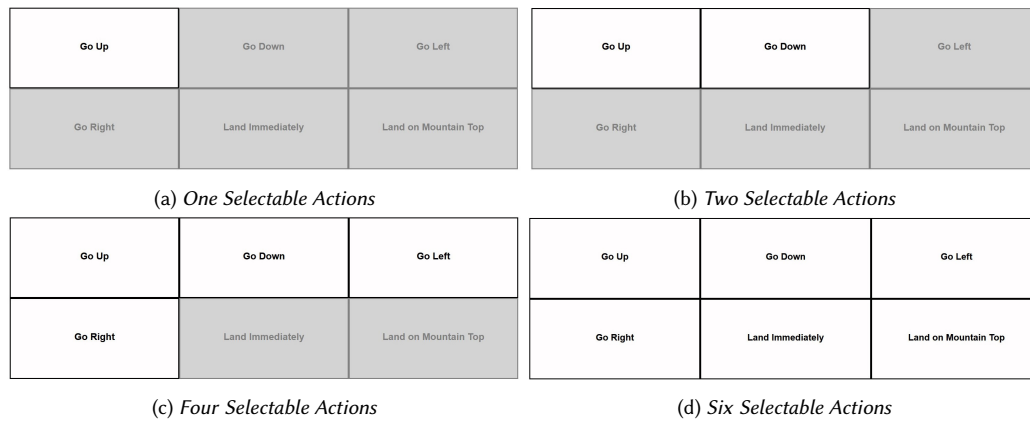


Fig. 1. Differences in the decision-making user interface depending on experimental conditions.

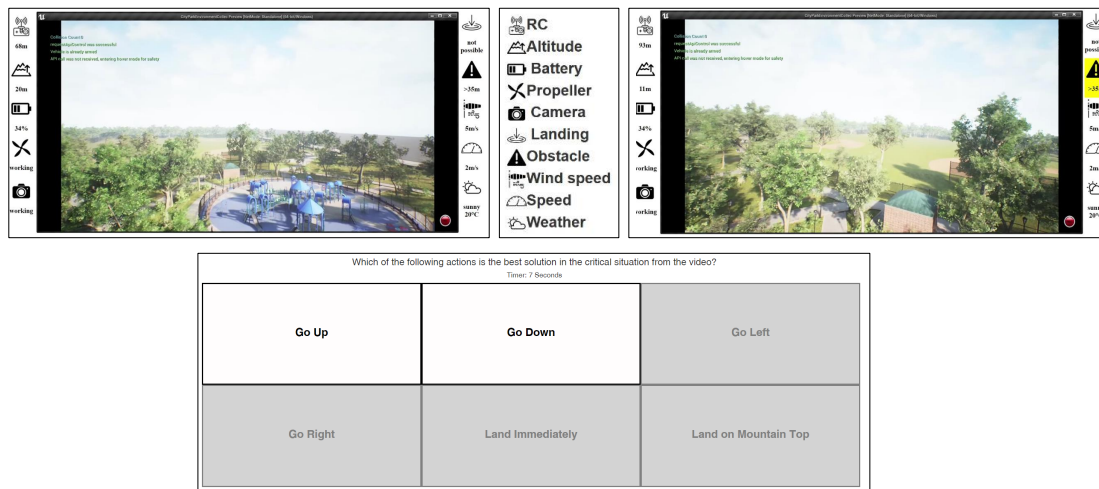


Fig. 2. Screenshots of the Drone Monitoring Interface. On the top left, a screenshot of the demo task is shown, where participants see a video recorded by the drone and information about the current status of the drone. In the top middle, the legend of all icons which were used in the interface is shown. This legend was shown to the participants during the entire video of the drone. Ten seconds before the participants had to make a decision, the drone entered a critical situation, which was indicated by an auditory signal and highlighted icons for critical values, as shown in the top right. After these ten seconds, the participants saw six possible actions, and depending on the experimental condition some of them were grayed out indicating they were not available as seen in the bottom screenshot of the *Two Selectable Actions* condition. In this demo task, the critical situation was the drone approaching trees which could be resolved by going up.

Figure 2 shows the monitoring interface and video feed of the drone. Sensor values included: the distance to the remote control, the current altitude and battery level, if the propeller and camera are working, if it is possible to land immediately, the distance to the next obstacle, the current wind speed and speed of the drone, and the current weather condition.

Toward the end of each video, the drone was in a critical situation that required a handover to the human overseer, which was indicated by an auditory signal. At this point, the video highlighted the sensor values related to the critical situation. After ten seconds a second auditory signal was played, the video disappeared and six actions were displayed (see Figure 2, bottom): fly higher, fly lower, fly to the left, fly to the right, land immediately, or land at a (varying) location. Depending on the condition some of the actions were grayed out (see below). Participants were instructed to choose an action as quickly as possible and had up to seven seconds to decide. To resolve the critical situation successfully, participants had to consider the corresponding sensor values and video feed to consider the possibility of landing at specific locations.

There were three types of critical situations: weather conditions, technical problems, and physical obstacles. Weather conditions caused a critical situation if the wind speed passed a certain threshold (unknown to participants) or if the current weather was rainy, snowy, windy, or foggy and the altitude was too high or too low. Technical problems occurred when the camera or propeller stopped working or the battery was too low. Physical obstacles led to critical situations when the distance to the next obstacle was too little or flying at the current altitude was dangerous.

After deciding on an action, participants were asked to fill out three open-ended questions to justify their answer (see supp. material). These ensured that people did not choose actions at random. If they did not choose an action, they were asked to give a reason. After that they were provided with feedback stating if their chosen action saved the drone or why it crashed.

After the 1st, 5th, 9th, and 10th (last) trial, participants filled out a questionnaire with Likert-scale items assessing their perceived autonomy, meaningfulness, task motivation, work satisfaction, and mental demand + temporal demand (which we combined to users' task load). In the final trial, there was no action that resolved the critical situation, always leading to a crash. This was to ensure that participants experienced at least one crash so that we could additionally ask questions related to their perceived responsibility in the last questionnaire. Table 1 shows the items in the questionnaire relevant to this paper. See Table 5 in the Appendix for the full set of questions asked, including additional exploratory variables, such as stress, trust, and responsibility-related variables. Since our analysis focused on the perceived autonomy and its impact on motivational variables, we restricted our results to the variables most relevant for these goals.

In the beginning of the experiment, participants received written instructions about the task, the meaning of sensor values, the types of critical situations, and possible actions. After the participants had all the required information, they were instructed to solve a trial task. In the trial task, they had the opportunity to make their own decision. Regardless of their decision, they got to know which action would have resolved the critical situation successfully. We also provided them with examples for this task on how to fill the open-ended questions. Exact instructions and screenshots from all steps of the experiment are provided in the supplementary material.

3.4 Implementation and Video material

The study we created was based on a previous study performed by Gundappa et al. [32]. They used their simulation framework that extends Microsoft AirSim [76] and can be used to create a variety of handover situations. In total, they created 20 different scenarios using two existing AirSim environments (i.e. park and valley environment). In our study, we used ten scenarios that had only one action that resolved the critical situation successfully. We made this choice so that there is room for wrong decisions in the *Two* and *Four* Conditions, while we did not want to make the participants lose trust in the recommendations by making correct actions not selectable. One of the scenarios is provided in the supplementary materials. We embedded the videos in a web page, which enabled us to create this interactive task and recruit participants on the Prolific Platform.

Table 1. An overview of the variables and corresponding questionnaire items collected and analyzed in this paper, the rounds after which each question was asked, and the sources from which they were adapted.

Concept	Question(s)	Ref
Perceived Autonomy	'I am so restricted by the available actions that I can hardly make my own decisions.' (<i>inversely coded</i>) 'I can make decisions about how to prevent crashes independently.'	[89]
Perceived Meaningfulness	'The work I do is important.' 'The work I do is meaningful.'	[77]
Task Motivation	'I have fun performing the task.'	[25]
Task Satisfaction	'I feel fulfilled performing the task.'	[65]
Task Load	'The task was mentally demanding.' 'The pace of the task was hurried or rushed.'	[35]

3.5 Measured Variables

We used self-report items to measure perceived *autonomy*, *meaningfulness*, *task motivation*, *task satisfaction* and *task load* on a 7-point Likert scale (1 - strongly disagree to 7 - strongly agree). The items we used are shown in Table 1 along with references to the literature from which we sourced these items. All additionally measured variables not included in this paper are shown in Table 5 in the Appendix. Self-report items were measured multiple times (after the 1., 5., 9., final scenario).

Note that given the substantial overall correlation between task motivation and task satisfaction ($r = .87$), for the remainder of this paper, we decided to only report task motivation because the findings for task satisfaction were nearly identical.

Additionally, for all ten scenarios, we measured two decision-related variables: *decision accuracy* and *decision time*. Decision accuracy was calculated for each participant as the percentage of selected actions that led to a successful outcome (i.e. drone not crashing) in the first nine rounds. If a participant did not select an action, this was counted as inaccurate. The last scenario was excluded because all actions led to a crash. Decision time was measured as the time after participants saw the available actions until they chose a specific action.

4 Results

4.1 Analysis

We decided to split our analysis into three parts, each requiring different approaches for statistical analysis. In subsection 4.2, we report the overall effects of *Selectable Actions* on our dependent variables, averaging results from all scenarios and repeated questionnaires. To test for significant differences, we chose a Kruskal-Wallis test combined with Dunn test with Bonferroni correction for the post-hoc comparisons of all experimental conditions. In subsection 4.3, we were interested in understanding how *Selectable Actions* impacted decision-related variables, perceived autonomy, and motivational factors over time. To this end, we treated *Decision Round* as an additional independent variable and report results from linear mixed models that test for interaction effects of the *Selectable Actions* and the *Decision Round*. Finally, in subsection 4.4, we combined the data of those conditions where participants had at least two actions to choose from (i.e. *Two*, *Four*, and *Six* (informed by our results in subsection 4.2)). Our goal was to analyze the effects of perceived autonomy, meaningfulness, and task motivation on decision accuracy and decision time. To this end, we

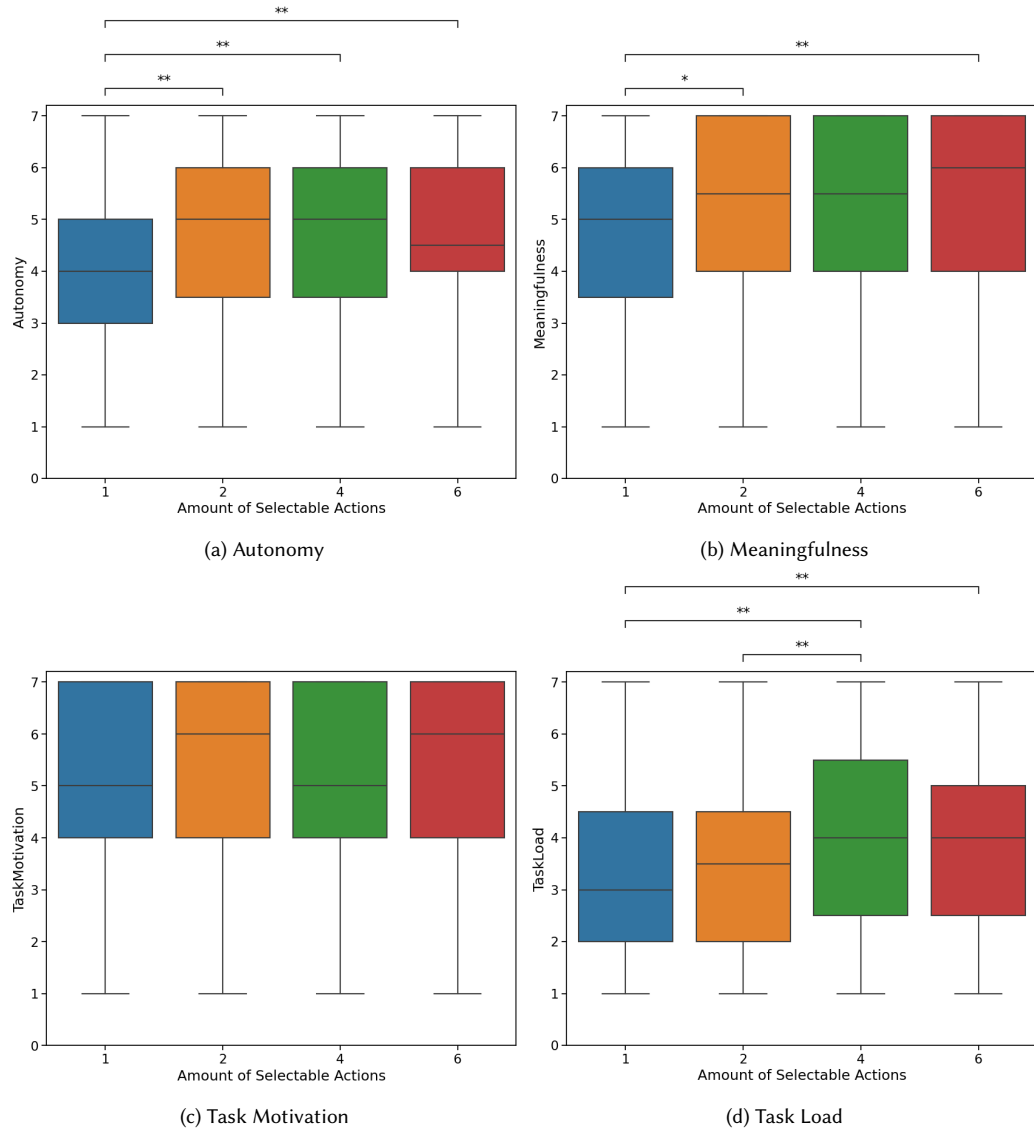


Fig. 3. Differences in perceived autonomy, meaningfulness, task motivation, and task load on a 7-point Likert Scale between the experimental conditions

report results from linear mixed models where perceived autonomy, meaningfulness, and motivation were added as predictors beyond *Selectable Actions* and the *Decision Round*.

4.2 The impact of the number of *Selectable Actions*

Research Questions 1a-f asked about the effects of *Selectable Actions* on various decision-related and self-report variables.

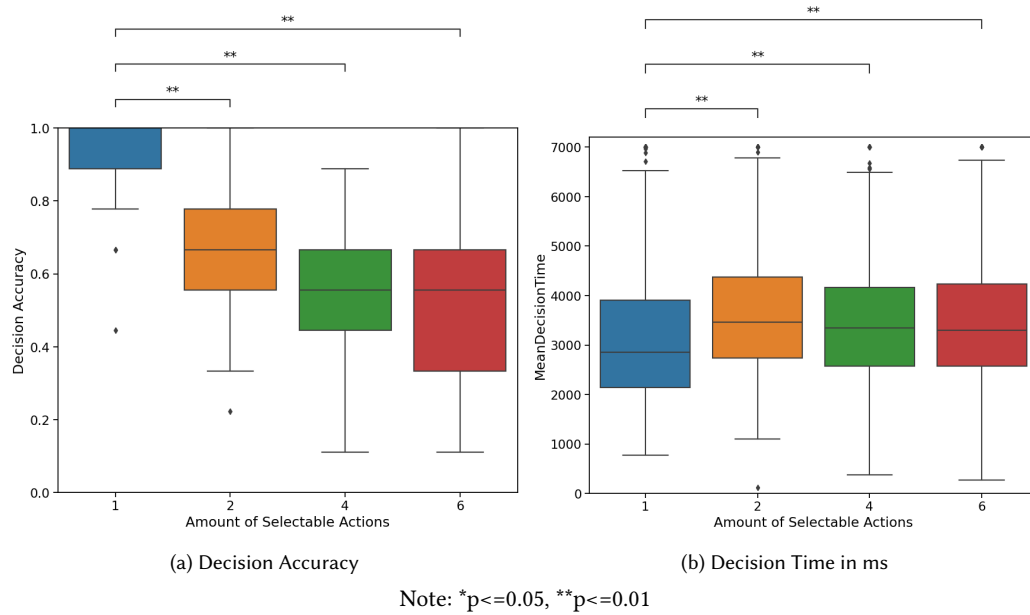


Fig. 4. Differences in Decision Accuracy in Percentage and Decision Time in milliseconds between the Experimental Conditions

Selectable Actions significantly affected participants' perceived autonomy ($H = 38.35$, $p < .001$), meaningfulness ($H = 15.52$, $p = .001$), and task load ($H = 30.71$, $p < .001$). As visualized by the box-plots in Figure 3, with *One Selectable Action*, participants felt less autonomous than having *Two* ($p < .001$), *Four* ($p < .001$), or *Six* ($p < .001$). Participants judged the task as less meaningful when having *One Selectable Action* compared to having *Two* ($p = .02$), or *Six* ($p = .001$). Also, with *One Selectable Action* they had a lower task load than with *Four* ($p < .001$), or *Six* ($p < .001$). In contrast, there was no significant difference for Motivation across all conditions (see Figure 3c).

The number of *Selectable Actions* also significantly affected the decision accuracy (Figure 4a) ($H = 128.88$, $p < .001$). Pairwise post-hoc comparisons showed participants having *One Selectable Action* to make better decisions than participants having *Two* ($p < .001$), *Four* ($p < .001$), or *Six* ($p < .001$) (Figure 4a). Although visually there is a trend that decision accuracy decreased with an increasing number of selectable actions, no significant differences were found when comparing conditions with more than one action to choose from.

We also found a significant difference in the *decision time* between the experimental conditions ($H = 28.79$, $p < .001$). Participants who had *One Selectable Action* decided faster than those with *Two* ($p < .001$), *Four* ($p < .001$), or *Six* ($p < .001$) (Figure 4b). No significant differences were found when comparing conditions with more than one action to choose from.

4.3 Changes in the effects of *Selectable Actions* over time

The findings up to now indicate that there are general differences between the experimental conditions but mostly between the condition with *One Selectable Action* and the conditions that had at least some available choice for an action. Given that the effects of varying *Selectable Actions* may only unfold over time or may grow stronger over time,

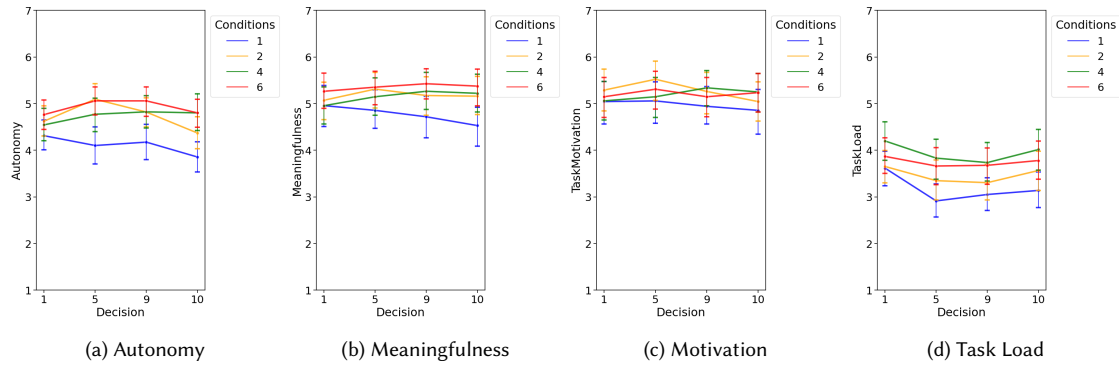


Fig. 5. Differences in perceived autonomy, meaningfulness, task motivation, and task load over time in each condition on a 7-point Likert Scale

we decided to additionally examine interactive effects of the *Selectable Actions* and the *Decision Round*, that is after how many scenarios the questionnaire was asked (1st, 5th, 9th, or 10th).

Figure 5 and Table 2 show the results of linear mixed model analyses with the dependent variables perceived autonomy, meaningfulness, motivation, task load, and decision time. The decision round is nested within participants as every participant experienced 10 decision rounds where we captured the relevant dependent variables four times. As predictors in the linear mixed models we included a) the *Selectable Actions* as a dummy-coded variable with *One Selectable Action* as the reference condition, b) the four *Decision Rounds* with Decision 1 as the reference condition, and c) the interaction effect between the *Selectable Actions* and the *Decision Round*.

For perceived autonomy, Figure 5a together with the significant interaction effects between the *Decision Round* and the *Selectable Actions* reported in Table 2 show that over time, the differences in perceived autonomy become larger. While at decision 1, there were comparably small differences for *Selectable Actions*, the differences between *One Selectable Action* and the other conditions were larger for later rounds.

For perceived meaningfulness, only for the ninth and for the final round were there significant interaction effects (see Table 2). This indicates that over time the difference in perceived meaningfulness between the condition with *One Selectable Action* and the other conditions became larger (see Figure 5b).

For task motivation (Figure 5c) and task load (Figure 5d) there were no clear effects other than a main effect for the *Decision Round* indicating a lower perceived task load for later rounds.

For the decision time, there were also no clear effects other than participants taking less time for the later rounds than for the first round.

Building on the results up to this point, we can respond to the Research Questions 1a-1f: The number of *Selectable Actions*...

RQ1a. **affected** users' perceived autonomy. The differences between *One Selectable Action* and the other conditions was stronger for later rounds than for *Decision Round 1*.

RQ1b. **affected** users' perceived meaningfulness. The differences between *One Selectable Action* and the other conditions was stronger for later rounds than for *Decision Round 1*.

RQ1c. **did not affect** users' task motivation.

Table 2. Linear mixed model analyses with the dependent variables perceived autonomy, meaningfulness, motivation, task load, and decision time

	<i>Dependent variable:</i>				
	Autonomy	Meaningfulness	TaskMotivation	TaskLoad	MeanDecisionTime
Intercept	4.31** (0.18)	4.96** (0.21)	5.04** (0.22)	3.29** (0.15)	4,011.33** (163.79)
2 Actions	0.31 (0.25)	0.12 (0.29)	0.25 (0.31)	0.09 (0.21)	473.12* (231.63)
4 Actions	0.23 (0.25)	−0.00 (0.30)	0.02 (0.31)	0.29 (0.21)	−97.10 (232.48)
6 Actions	0.45 (0.25)	0.31 (0.30)	0.10 (0.31)	0.28 (0.21)	56.87 (232.48)
Decision 5	−0.21 (0.15)	−0.10 (0.13)	0.01 (0.17)	−0.23** (0.05)	−1,101.45** (210.27)
Decision 9	−0.14 (0.15)	−0.24 (0.13)	−0.10 (0.17)	−0.19** (0.05)	−1,141.52** (210.27)
Decision 10	−0.46** (0.15)	−0.43** (0.13)	−0.19 (0.17)	−0.16** (0.05)	−931.03** (210.27)
2 Actions:Decision 5	0.68** (0.21)	0.34 (0.18)	0.22 (0.24)	0.13 (0.07)	−157.15 (297.36)
4 Actions:Decision 5	0.44* (0.22)	0.29 (0.18)	0.07 (0.24)	0.11 (0.07)	417.74 (298.45)
6 Actions:Decision 5	0.50* (0.22)	0.19 (0.18)	0.15 (0.24)	0.17* (0.07)	289.13 (298.45)
2 Actions:Decision 9	0.33 (0.21)	0.34 (0.18)	0.07 (0.24)	0.07 (0.07)	287.27 (297.36)
4 Actions:Decision 9	0.42 (0.22)	0.55** (0.18)	0.38 (0.24)	0.03 (0.07)	706.60* (298.45)
6 Actions:Decision 9	0.43* (0.22)	0.40* (0.18)	0.10 (0.24)	0.12 (0.07)	632.72* (298.45)
2 Actions:Decision 10	0.20 (0.21)	0.51** (0.18)	−0.06 (0.24)	0.13 (0.07)	−46.01 (297.36)
4 Actions:Decision 10	0.71** (0.22)	0.69** (0.18)	0.38 (0.24)	0.10 (0.07)	447.12 (298.45)
6 Actions:Decision 10	0.49* (0.22)	0.54** (0.18)	0.28 (0.24)	0.13 (0.07)	398.21 (298.45)
Observations	1,096	1,096	1,096	1,096	1,096

Note: The reference condition for Actions is 1 Action; the reference condition for Decision is Decision 1. *p<0.05; **p<0.01

- RQ1d. **affected** users' task load. There was no clear significant interaction effects between the number of *Selectable Actions* and the *Decision Round*.
- RQ1e. **affected** the decision accuracy in the drone handover task. Participants who had *One Selectable Action* performed better than participants in the other conditions.
- RQ1f. **affected** users' decision time. Participants who had *One Selectable Action* took less time to decide than participants in the other conditions. There was no clear significant interaction effects between the number of *Selectable Actions* and the *Decision Round*.

Table 3. Linear mixed models with the dependent variable decision accuracy

	<i>Dependent variable:</i>			
	PercentageCorrectAnswers			
	(1)	(2)	(3)	(4)
Intercept	0.25** (0.06)	0.54** (0.06)	0.48** (0.06)	0.55** (0.07)
Autonomy	0.05** (0.01)			
Meaningfulness		−0.01 (0.01)		
TaskMotivation			−0.00 (0.01)	
TaskLoad				−0.02 (0.01)
4 Actions	−0.02 (0.06)	−0.02 (0.06)	−0.02 (0.06)	−0.02 (0.06)
6 Actions	−0.16** (0.06)	−0.15** (0.06)	−0.15** (0.06)	−0.15* (0.06)
Decision 5	0.20** (0.05)	0.22** (0.05)	0.22** (0.05)	0.22** (0.05)
Decision 9	0.13** (0.05)	0.14** (0.05)	0.14** (0.05)	0.14** (0.05)
4 Actions:Decision 5	−0.10 (0.07)	−0.11 (0.07)	−0.11 (0.07)	−0.11 (0.07)
6 Actions:Decision 5	−0.04 (0.07)	−0.05 (0.07)	−0.05 (0.07)	−0.05 (0.07)
4 Actions:Decision 9	−0.00 (0.07)	0.00 (0.07)	0.00 (0.07)	0.00 (0.07)
6 Actions:Decision 9	0.13 (0.07)	0.14* (0.07)	0.14* (0.07)	0.14* (0.07)
Observations	615	615	615	615

Note: The reference condition for Actions is 2 Actions; the reference condition for Decision is Decision 1. (1)-(4) reflect that those are four different linear mixed models, (1) tests the effect of perceived autonomy, (2) meaningfulness, (3) task motivation, and (4) task load on the percentage of correct answers. *p<0.05; **p<0.01

Table 4. Linear mixed models with the dependent variable decision time

	<i>Dependent variable:</i>			
	MeanDecisionTime			
	(1)	(2)	(3)	(4)
Intercept	4,579.66** (231.09)	4,119.91** (224.38)	3,937.71** (219.45)	4,646.55** (223.84)
Autonomy	−20.59 (36.06)	-	-	-
Meaningfulness	-	71.87* (31.19)	-	-
TaskMotivation	-	-	103.36** (28.74)	-
TaskLoad	-	-	-	−47.87 (46.26)
4 Actions	−571.84* (227.14)	−561.84* (225.88)	−546.34* (224.71)	−560.91* (227.10)
6 Actions	−413.33 (227.18)	−430.06 (225.93)	−401.48 (224.65)	−407.41 (227.08)
Decision 5	−1,248.90** (209.45)	−1,275.79** (209.56)	−1,282.57** (209.07)	−1,263.46** (208.92)
Decision 9	−850.22** (208.88)	−861.54** (209.45)	−851.25** (208.97)	−859.80** (208.94)
Decision 10	−982.27** (208.96)	−983.29** (209.44)	−951.58** (209.09)	−978.43** (208.87)
4 Actions:Decision 5	569.88 (296.45)	578.33 (297.26)	589.73* (296.64)	573.88 (296.47)
6 Actions:Decision 5	442.64 (296.39)	457.13 (297.30)	453.53 (296.62)	447.85 (296.47)
4 Actions:Decision 9	421.05 (296.34)	404.43 (297.33)	387.45 (296.74)	417.49 (296.47)
6 Actions:Decision 9	347.48 (296.34)	341.12 (297.26)	342.46 (296.61)	347.95 (296.48)
4 Actions:Decision 10	503.66 (296.89)	480.36 (297.31)	447.91 (296.88)	491.59 (296.47)
6 Actions:Decision 10	450.20 (296.51)	442.54 (297.26)	409.64 (296.77)	444.20 (296.47)
Observations	820	820	820	820

Note: The reference condition for Actions is 2 Actions; the reference condition for Decision is Decision 1. (1)-(4) reflect that those are four different linear mixed models, (1) tests the effect of perceived autonomy, (2) meaningfulness, (3) task motivation, and (4) task load on the percentage of correct answers. *p<0.05; **p<0.01

4.4 Motivational variables and their relation to decision-making

Research Questions 2a-b asked whether perceived autonomy, meaningfulness and task motivation affect decision-related variables. To explore this, we calculated another set of linear mixed models with the dependent variables decision accuracy (Table 3) and decision time (Table 4). For these analyses, we excluded participants from condition *One*, since there was virtually no variance between participants' decision accuracy in this condition given that they always received the successful solution as a recommendation by the AI-based system. We also excluded the final round of the study in which all participants were forced to take a decision leading to a crash, thus having no variance between participants in their decision accuracy.

The linear mixed models thus included dummy-coded variables that modeled *Two Selectable Action* as the reference condition, dummy-coded variables that modeled *Decision Round 1* as the reference condition, and interaction effects between *Selectable Actions* and *Decision Round*. We included autonomy, meaningfulness, and task motivation in different linear mixed models to test whether they affect the decision-related variables beyond the experimental condition, beyond the decision round, and beyond the interactive effects between the experimental condition and the decision round.

For decision accuracy, we found that perceived autonomy was significantly positively related to participants' decision accuracy. This was not the case for perceived meaningfulness and motivation.

For the decision time, we found that autonomy was not associated with decision time. Instead, we found that higher motivation and meaningfulness were associated with significantly longer decision times.

Regarding Research Questions 2, we can thus respond that,

RQ2a. the accuracy of decisions made in the drone handover task was **positively related** to perceived autonomy.

RQ2b. participants' decision time was **positively related** to their perceived task meaningfulness and motivation.

5 Discussion

The goal of this study was to examine whether an AI-based system that restricts the available choices of users to different degrees affects decision-related variables, perceived autonomy, and motivational aspects of human-AI interaction. The main findings of this study are

- (1) that having a choice positively affected perceived autonomy and perceived meaningfulness of the task with effects growing stronger over time,
- (2) that when having a choice, perceived autonomy was associated with higher decision accuracy, while meaningfulness and motivation were associated with longer decision times, and
- (3) taking the choice away from the participants positively affected their decision accuracy, owing to the nature of our task and the highly accurate AI-based support.

Together these findings provide first evidence on the impact that AI-based systems can have on workers psychological needs and motivation and how this might affect overall work performance. In the following, we discuss these results in more detail and point out limitations of our study and opportunities for future work.

5.1 The effects of having a choice

Restricting choices for our participants improved their decision accuracy, decreased their decision time, and reduced the task load. These findings are in line with research showing that highly accurate and reliable AI-based systems can improve joint human-AI decision accuracy [22]. The positive effect on decision accuracy is also in line with De

Vreede et al.'s [18] finding that restricting users' choices can positively affect joint human-AI decision accuracy. This is positive news for the use of AI systems in decision-making tasks. In our study, this effect was most clearly reflected in the better decision accuracy of the participants who only had one available action compared to the other conditions. Eventually, these participants had no real "choice" than to approve the system recommendation or to wait and ignore the recommendation, which would have led to a drone crash. At first glance, this may sound like an unrealistic scenario – why even have a human overseer in such situations? However, given emerging legal regulations, such as the European Union's AI act, demanding the involvement of human overseers for high-risk tasks [27, 79] (which includes the operation of autonomous drones, cf. Art. 6 and Annex 1(20) of the AI Act), we believe the scenario of having a human overseer simply approving AI-based recommendations with a limited degree of choice is realistic, though legal certainty on this matter will only be reached in the future. Importantly, those positive effects of restricting the *Selectable Actions* on decision accuracy strongly depend on the reliability of the AI-based recommendations given that it is to expect that overseers will tend to follow also inaccurate AI-based recommendations [27, 62]. Thus, the positive effects of restricting people's choices through AI-based recommendations will be less clear in cases where the AI-based system does not always offer the correct solution. Given that perfect and near-perfect recommendations by AI-based systems are unusual and will probably also be unusual in the foreseeable future, this is an important consideration when restricting users' autonomy through AI-based systems.

The positive effects of restricting participants' choices on decision accuracy stand in contrast to their motivational effects. In contrast to only having one selectable action, having even just some choice (i.e., *Two Selectable Actions*) positively affected participants' perceived autonomy and meaningfulness of their task. Crucially, the differences between *One* and other conditions became stronger over time. Consistent with the Self-Determination theory [73], the work design literature [34, 57], and the SMART model of work design [68], offering users some choices in their interaction with an AI-based system can be positively related to their need for autonomy and can have beneficial motivational effects reflected in their perceived meaningfulness of the task. In contrast, when users have little choice than to follow the recommendation of an AI-based system, over time they may start feeling useless as their behavior in the task is void of perceived meaning. In the long term, ignoring such detrimental effects on human needs and motivational variables may not only be negative for users motivation but also for their well-being with possible negative effects of jobs with little motivating appeal on people's mental health [66].

Consistent with this interpretation, whereas the number of selectable choices did not affect participants' motivation (and task satisfaction) in this study, we may expect that such effects could show during long term interaction with the system. Based on the SMART model and the broader work design literature, perceived meaningfulness is one of the strongest predictors of worker motivation and satisfaction over time [40, 66, 68]. In our study, participants only experienced 10 decision rounds and an effective interaction with a system for about 10 minutes. This may not be enough to find long-term effects on user motivation. But even with just 10 minutes of interacting with the system, our participants started to question the meaningfulness of their role in this task. Thus, one key implication of the current study is that it will be crucial to conduct more long-term interaction studies to test for motivational effects in human-AI collaboration – studies that are still rare in human-AI research [21]. Additionally, the effects on perceived autonomy and meaningfulness are important insights for the future design of human oversight jobs and of jobs where AI-based systems affect users' autonomy.

What may also be important to consider for future research is to find the right balance between considering users' need for autonomy and overwhelming users with too many choices. De Vreede et al. [18] showed that too much autonomy i.e., letting people decide freely when to use and when not to use an AI-based system, can be detrimental for

joint human-AI decision accuracy. In our study, having *Four* and *Six* available choices led to higher perceived task load and both conditions were related to lower decision accuracy than having one available option. Eventually, having too much choice could even undermine users' perceived autonomy in a way that users will feel that they are not in a good position to make informed decisions. This curvilinear effect of autonomy is also reflected in evidence from the work design literature that shows that it is important to find a suitable degree of autonomy because autonomy misfit (i.e., too low and too much autonomy) can undermine people's motivation and well-being at work [81], and can lead to detrimental effects such as users manipulating input data for AI-based systems to regain their autonomy [82].

In that regard, one implication of our study could be that as long as users have some choice, increasing available choices even more may not have further beneficial effects on variables such as perceived autonomy or meaningfulness. To put it differently, a certain degree of restriction in a way that an AI-based system restricts ineffective actions may not have detrimental effects on motivational variables important for effective long-term human-AI collaboration. In safety critical scenarios where it is important to minimize erroneous decisions, this is positive news. It may be possible to restrict users' choices in human-AI interactions in a way that leads to good joint performance without detrimentally impacting users' perceived autonomy — as long as they are left with some choice. Clearly, this conclusion comes with ethical issues. For example, knowing that it is enough to ensure that people have some choice in order to uphold their perceived autonomy means that it may be enough to provide users with pseudo-choices or with the ability to slightly change the system with effects that are of little practical importance [20]. Responsible design of human-AI collaboration needs to consider such potentially unethical ways of making people feel autonomous when all they have are pseudo-choices.

5.2 Perceived autonomy is positively related to decision accuracy - meaningfulness and motivation are positively related to decision time

In the conditions with some choice, those participants who felt more autonomous were also the ones who performed better. This supports propositions and findings by the work design literature that perceived autonomy can contribute to task performance [40, 68]. In theory this kind of a positive effect may have been expected to also correspond to a stronger motivation of users. Yet, motivation was not significantly positively associated with decision accuracy in our analyses. We propose that this again calls for future research that examines the effects of design decisions that affect motivational aspects of human-AI collaboration in the long term. Potentially, in longer-term interactions it will be possible to provide insights on the expected relations between perceived autonomy, meaningfulness, and task motivation.

Interestingly, we found that task meaningfulness and motivation was related to a longer time that participants took for their decisions. On the one hand, this could mean that those participants who found the task to be more meaningful and who were motivated did their best in the task trying to think through their decisions for the *Selectable Actions*. On the other hand, this association could mean that those people who found the task less meaningful and were less motivated wanted to finish it more quickly. Both of these interpretations for the findings would have been stronger if we had found relations between meaningfulness and/or motivation and decision accuracy that went beyond the effect of the experimental manipulations. Future research needs to more closely examine the relation between motivational variables and decision-related variables to uncover relations that are likely reciprocal: motivational variables affect decision-related variables (e.g., high motivation leads to high decision accuracy), and decision-related variables affect motivational variables (e.g., high decision accuracy leads to high motivation) [85].

5.3 Limitations and Future Work

There are three main limitations of the current study that we want to highlight. First, participants only experienced a simulated drone oversight task without any real-world consequences. This can affect self-report items relating to users' perception of the oversight task. Nevertheless, we tried to simulate a realistic situation by showing participants videos of drone flights and by simulating time pressure.

Second, participants only experienced ten scenarios and all of them ended in critical situations. In reality, people oversee drones for a longer period of time and optimally experience only a low number of critical situations. This may also be one reason for not finding significant differences between the experimental conditions for variables such as task motivation that may only play out after a prolonged period of time. Thus, for more generalizable insights on the role of selectable options on long-term implications on variables important for job satisfaction, future work would optimally simulate longer-term interactions with AI-based systems. However, this comes with many additional complexities and costs. We believe that for future research it will be worth investing these efforts given the findings and implications of the current study and given other recent research that highlights the importance of understanding the impact of human-AI interaction design on motivational variables and human needs [6].

Third, participants presumably had limited experience in flying drones and were no experts in human oversight tasks. This lack of expertise might have affected their overall perceived autonomy, as well as further motivational variables that we captured in this study. Still, we believe that the current study included a task where our participants were able to immerse themselves into and where they were capable of quickly understanding their responsibilities in the task. We are not sure how conducting the current study with experts would change the current results. This may depend on whether the person tasked with overseeing the drone was an expert drone pilot or an expert overseer. For expert drone pilots, reducing the autonomy may be more impactful than for expert overseers who may be used to having only limited points in time where they are supposed to intervene. However, this is a hypothesis for future work where it may be interesting to more closely examine motivational differences between people with different backgrounds who are supposed to fill the role of human overseers in high-risk tasks.

6 Conclusion

Beyond optimizing human-AI interactions for cognitive variables, it is important to also consider motivational variables which are often associated with satisfying human psychological needs [68, 73]. In this regard, the current study examined the impact of restricting people's choices on their perceived autonomy and on downstream motivational variables and task performance via an AI-based decision support in an oversight task. Our findings highlight that restricting choices can be effective in helping people make accurate decisions. At the same time, restricting choices can undermine human overseers' perceived autonomy and important motivational variables such as the perceived meaningfulness of a task, particularly in the long term. The finding that people who felt more autonomous were also the ones who showed better decision accuracy further highlights the potential positive effects that considering human needs in human-AI interaction design can have on the joint performance of humans and AI-based systems. The most promising direction for future research that we see based on the current study is to conduct studies where humans and systems interact over the long term because this will help to better understand the impact of motivational variables in human-AI interactions that may only become stronger over time.

Acknowledgments

This work was funded by DFG grant 389792660 as part of TRR 248 – CPEC, see <https://perspicuous-computing.science>. We thank Emilia Ellsiepen for all the support regarding the study design and execution.

References

- [1] Yasmeen Alufaisan, Laura R. Marusich, Jonathan Z. Bakdash, Yan Zhou, and Murat Kantarcioglu. 2021. Does Explainable Artificial Intelligence Improve Human Decision-Making? *Proceedings of the AAAI Conference on Artificial Intelligence* 35, 8 (May 2021), 6618–6626. <https://doi.org/10.1609/aaai.v35i8.16819> Number: 8.
- [2] Daniel L. Ames and Susan T. Fiske. 2013. Intentional Harms Are Worse, Even When They're Not. *Psychological Science* 24, 9 (Sept. 2013), 1755–1762. <https://doi.org/10.1177/0956797613480507> Publisher: SAGE Publications Inc.
- [3] Ksenia Appelganc, Tobias Rieger, Eileen Roesler, and Dietrich Manzey. 2022. How Much Reliability Is Enough? A Context-Specific View on Human Interaction With (Artificial) Agents From Different Perspectives. *Journal of Cognitive Engineering and Decision Making* 16, 4 (Dec. 2022), 207–221. <https://doi.org/10.1177/15553434221104615> Publisher: SAGE Publications.
- [4] J. Elin Bahner, Anke-Dorothea Hüper, and Dietrich Manzey. 2008. Misuse of automated decision aids: Complacency, automation bias and the impact of training experience. *International Journal of Human-Computer Studies* 66, 9 (Sept. 2008), 688–699. <https://doi.org/10.1016/j.ijhcs.2008.06.001>
- [5] Gagan Bansal, Tongshuang Wu, Joyce Zhou, Raymond Fok, Besmira Nushi, Ece Kamar, Marco Tulio Ribeiro, and Daniel Weld. 2021. Does the Whole Exceed its Parts? The Effect of AI Explanations on Complementary Team Performance. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems (CHI '21)*. Association for Computing Machinery, New York, NY, USA, 1–16. <https://doi.org/10.1145/3411764.3445717>
- [6] Zana Bućinca. 2024. Optimizing Decision-Maker's Intrinsic Motivation for Effective Human-AI Decision-Making. In *Extended Abstracts of the 2024 CHI Conference on Human Factors in Computing Systems (CHI EA '24)*. Association for Computing Machinery, New York, NY, USA, Article 439, 5 pages. <https://doi.org/10.1145/3613905.3638179>
- [7] John Buchanan and Ned Kock. 2001. Information Overload: A Decision Making Perspective. In *Multiple Criteria Decision Making in the New Millennium*, G. Fandel, W. Trockel, C. D. Aliprantis, Dan Kovenock, Murat Köksalan, and Stanley Zionts (Eds.). Vol. 507. Springer Berlin Heidelberg, Berlin, Heidelberg, 49–58. https://doi.org/10.1007/978-3-642-56680-6_4 Series Title: Lecture Notes in Economics and Mathematical Systems.
- [8] Zana Bućinca, Siddharth Swaroop, Amanda E Paluch, Susan A Murphy, and Krzysztof Z. Gajos. 2024. Towards Optimizing Human-Centric Objectives in AI-Assisted Decision-Making With Offline Reinforcement Learning. *arXiv preprint arXiv:2403.05911* (2024).
- [9] Zana Bućinca, Phoebe Lin, Krzysztof Z. Gajos, and Elena L. Glassman. 2020. Proxy tasks and subjective measures can be misleading in evaluating explainable AI systems. In *Proceedings of the 25th International Conference on Intelligent User Interfaces*. ACM, Cagliari Italy, 454–464. <https://doi.org/10.1145/3377325.3377498>
- [10] Zana Bućinca, Maja Barbara Malaya, and Krzysztof Z. Gajos. 2021. To Trust or to Think: Cognitive Forcing Functions Can Reduce Overreliance on AI in AI-assisted Decision-making. *Proceedings of the ACM on Human-Computer Interaction* 5, CSCW1 (April 2021), 1–21. <https://doi.org/10.1145/3449287>
- [11] Federico Cabitza, Andrea Campagner, Chiara Natali, Enea Parimbelli, Luca Ronzio, and Matteo Cameli. 2023. Painting the Black Box White: Experimental Findings from Applying XAI to an ECG Reading Setting. *Machine Learning and Knowledge Extraction* 5, 1 (March 2023), 269–286. <https://doi.org/10.3390/make5010017> Number: 1 Publisher: Multidisciplinary Digital Publishing Institute.
- [12] Francisco Maria Calisto, Carlos Santiago, Nuno Nunes, and Jacinto C. Nascimento. 2022. BreastScreening-AI: Evaluating medical intelligent agents for human-AI interactions. *Artificial Intelligence in Medicine* 127 (May 2022), 102285. <https://doi.org/10.1016/j.artmed.2022.102285>
- [13] J.P. Campbell, R.A. McCloy, S.H. Oppler, and C.E. Sager. 1993. A Theory of Performance. In *Personnel Selection in Organizations*, N. Schmitt and W.C. Borman (Eds.). Jossey-Bass Publishers, San Francisco, 35–70.
- [14] Federico Maria Cau, Hanna Hauptmann, Lucio Davide Spano, and Nava Tintarev. 2023. Supporting High-Uncertainty Decisions through AI and Logic-Style Explanations. In *Proceedings of the 28th International Conference on Intelligent User Interfaces (IUI '23)*. Association for Computing Machinery, New York, NY, USA, 251–263. <https://doi.org/10.1145/3581641.3584080>
- [15] Christopher P. Cerasoli, Jessica M. Nicklin, and Alexander S. Nassreelgawi. 2016. Performance, incentives, and needs for autonomy, competence, and relatedness: a meta-analysis. *Motivation and Emotion* 40, 6 (Sept. 2016), 781–813. <https://doi.org/10.1007/s11031-016-9578-2>
- [16] Alexander Chernev, Ulf Böckenholt, and Joseph Goodman. 2015. Choice overload: A conceptual review and meta-analysis. *Journal of Consumer Psychology* 25, 2 (2015), 333–358. <https://doi.org/10.1016/j.jcps.2014.08.002> eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1016/j.jcps.2014.08.002>
- [17] Anwesha Das, Zekun Wu, Iza Skrjanec, and Anna Maria Feit. 2024. Shifting Focus with HCEye: Exploring the Dynamics of Visual Highlighting and Cognitive Load on User Attention and Saliency Prediction. *Proceedings of the ACM on Human-Computer Interaction* 8, ETRA (May 2024), 1–18. <https://doi.org/10.1145/3655610>
- [18] Triparna De Vreede, Mukhunth Raghavan, and Gert-Jan De Vreede. 2021. Design Foundations for AI Assisted Decision Making: A Self Determination Theory Approach. In *Proceedings of the 54th Annual Hawaii International Conference on System Sciences*. <https://doi.org/10.24251/HICSS.2021.019>
- [19] Evangelia Demerouti and Arnold B Bakker. 2008. The Oldenburg Burnout Inventory: A good alternative to measure burnout and engagement. *Handbook of stress and burnout in health care* 65, 7 (2008), 1–25.
- [20] Berkeley J. Dietvorst, Joseph P. Simmons, and Cade Massey. 2018. Overcoming Algorithm Aversion: People Will Use Imperfect Algorithms If They Can (Even Slightly) Modify Them. *Management Science* 64, 3 (March 2018), 1155–1170. <https://doi.org/10.1287/mnsc.2016.2643>

- [21] Mengchen Dong, Jean-Francois Bonnefon, and Iyad Rahwan. 2023. Toward Human-Centered AI Management: Methodological Challenges and Future Directions. *SSRN Electronic Journal* (2023). <https://doi.org/10.2139/ssrn.4444322>
- [22] Simon Eisbach, Markus Langer, and Guido Hertel. 2023. Optimizing human-AI collaboration: Effects of motivation and accuracy information in AI-supported decision-making. *Computers in Human Behavior: Artificial Humans* 1, 2 (Aug. 2023), 100015. <https://doi.org/10.1016/j.chbah.2023.100015>
- [23] Franz Faul, Edgar Erdfelder, Albert-Georg Lang, and Axel Buchner. 2007. G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods* 39, 2 (May 2007), 175–191. <https://doi.org/10.3758/BF03193146>
- [24] Anna Maria Feit, Lukas Vordemann, Seonwook Park, Caterina Berube, and Otmar Hilliges. 2020. Detecting relevance during decision-making from eye movements for UI adaptation. *Eye Tracking Research and Applications Symposium (ETRA)* (2 2020). <https://doi.org/10.1145/3379155.3391321>
- [25] Marylène Gagné, Jacques Forest, Marie-Hélène Gilbert, Caroline Aubé, Estelle Morin, and Angela Malorni. 2010. The Motivation at Work Scale: Validation Evidence in Two Languages. *Educational and Psychological Measurement* 70, 4 (Aug. 2010), 628–646. <https://doi.org/10.1177/0013164409355698>
- [26] Jeannine A. Gailley and R. Frank Falk. 2008. Attribution of Responsibility as a Multidimensional Concept. *Sociological Spectrum* 28, 6 (Sept. 2008), 659–680. <https://doi.org/10.1080/02732170802342958> Publisher: Routledge _eprint: <https://doi.org/10.1080/02732170802342958>
- [27] Ben Green. 2022. The flaws of policies requiring human oversight of government algorithms. *Computer Law & Security Review* 45 (July 2022), 105681. <https://doi.org/10.1016/j.clsr.2022.105681>
- [28] Ben Green and Yiling Chen. 2019. Disparate Interactions: An Algorithm-in-the-Loop Analysis of Fairness in Risk Assessments. In *Proceedings of the Conference on Fairness, Accountability, and Transparency (FAT* '19)*. ACM. <https://doi.org/10.1145/3287560.3287563>
- [29] Ben Green and Yiling Chen. 2019. The Principles and Limits of Algorithm-in-the-Loop Decision Making. *Proc. ACM Hum.-Comput. Interact.* 3, CSCW (Nov. 2019), 50:1–50:24. <https://doi.org/10.1145/3359152>
- [30] Nina Grgić-Hlača, Christoph Engel, and Krishna P. Gummadi. 2019. Human Decision Making with Machine Assistance: An Experiment on Bailing and Jailng. *Proceedings of the ACM on Human-Computer Interaction* 3, CSCW (Nov. 2019), 1–25. <https://doi.org/10.1145/3359280>
- [31] Hongyan Gu, Yuan Liang, Yifan Xu, Christopher Kazu Williams, Shino Magaki, Negar Khanlou, Harry Vinters, Zesheng Chen, Shuo Ni, Chunxu Yang, Wenzhong Yan, Xinhai Robert Zhang, Yang Li, Mohammad Haeri, and Xiang ‘Anthony’ Chen. 2023. Improving Workflow Integration with xPath: Design and Evaluation of a Human-AI Diagnosis System in Pathology. *ACM Trans. Comput.-Hum. Interact.* 30, 2 (March 2023), 28:1–28:37. <https://doi.org/10.1145/3577011>
- [32] Ashwini Gundappa, Emilia Ellsiepen, Lukas Schmitz, Frederik Wiehr, and Vera Demberg. 2024. The Design of Informative Take-Over Requests for Semi-Autonomous Cyber-Physical Systems: Combining Spoken Language and Visual Icons in a Drone-Controller Setting. *arXiv:2409.08253* [cs.HC] <https://arxiv.org/abs/2409.08253>
- [33] J. Richard Hackman. 1980. Work redesign and motivation. *Professional Psychology* 11, 3 (June 1980), 445–455. <https://doi.org/10.1037/0735-7028.11.3.445>
- [34] J. Richard Hackman and Greg R. Oldham. 1975. Development of the Job Diagnostic Survey. *Journal of Applied Psychology* 60, 2 (April 1975), 159–170. <https://doi.org/10.1037/h0076546>
- [35] Sandra G. Hart and Lowell E. Staveland. 1988. Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research. In *Advances in Psychology*, Peter A. Hancock and Najmedin Meshkati (Eds.). Human Mental Workload, Vol. 52. North-Holland, 139–183. [https://doi.org/10.1016/S0166-4115\(08\)62386-9](https://doi.org/10.1016/S0166-4115(08)62386-9)
- [36] Lukas-Valentin Herm. 2023. IMPACT OF EXPLAINABLE AI ON COGNITIVE LOAD: INSIGHTS FROM AN EMPIRICAL STUDY. *ECIS 2023 Research Papers* (2023).
- [37] Pamela Hinds, Teresa Roberts, and Hank Jones. 2004. Whose Job Is It Anyway? A Study of Human-Robot Interaction in a Collaborative Task. *Human-Computer Interaction* 19, 1 (June 2004), 151–181. https://doi.org/10.1207/s15327051hci1901&2_7
- [38] Robert R. Hoffman, Shane T. Mueller, Gary Klein, and Jordan Litman. 2019. Metrics for Explainable AI: Challenges and Prospects. <https://doi.org/10.48550/arXiv.1812.04608> arXiv:1812.04608 [cs].
- [39] Antoine Hudon, Théophile Demazure, Alexander Karran, Pierre-Majorique Léger, and Sylvain Sénécal. 2021. Explainable Artificial Intelligence (XAI): How the Visualization of AI Predictions Affects User Cognitive Load and Confidence. In *Information Systems and Neuroscience*, Fred D. Davis, René Riedl, Jan vom Brocke, Pierre-Majorique Léger, Adriane B. Randolph, and Gernot Müller-Putz (Eds.). Springer International Publishing, Cham, 237–246. https://doi.org/10.1007/978-3-030-88900-5_27
- [40] Stephen E. Humphrey, Jennifer D. Nahrgang, and Frederick P. Morgeson. 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 (2007), 1332–1356. <https://doi.org/10.1037/0021-9010.92.5.1332>
- [41] Maia Jacobs, Melanie F. Pradier, Thomas H. McCoy, Roy H. Perlis, Finale Doshi-Velez, and Krzysztof Z. Gajos. 2021. How machine-learning recommendations influence clinician treatment selections: the example of antidepressant selection. *Translational Psychiatry* 11, 1 (Feb. 2021), 1–9. <https://doi.org/10.1038/s41398-021-01224-x> Publisher: Nature Publishing Group.
- [42] Patricia K. Kahr, Gerrit Rooks, Martijn C. Willemsen, and Chris C.P. Snijders. 2023. It Seems Smart, but It Acts Stupid: Development of Trust in AI Advice in a Repeated Legal Decision-Making Task. In *Proceedings of the 28th International Conference on Intelligent User Interfaces*. ACM, Sydney NSW Australia, 528–539. <https://doi.org/10.1145/3581641.3584058>
- [43] Moritz Körber. 2019. Theoretical Considerations and Development of a Questionnaire to Measure Trust in Automation. In *Proceedings of the 20th Congress of the International Ergonomics Association (IEA 2018)*, Sebastiano Bagnara, Riccardo Tartaglia, Sara Albolino, Thomas Alexander, and Yushi

- Fujita (Eds.). Springer International Publishing, Cham, 13–30. https://doi.org/10.1007/978-3-319-96074-6_2
- [44] Vivian Lai, Chacha Chen, Alison Smith-Renner, Q. Vera Liao, and Chenhao Tan. 2023. Towards a Science of Human-AI Decision Making: An Overview of Design Space in Empirical Human-Subject Studies. In *2023 ACM Conference on Fairness, Accountability, and Transparency (FAccT '23)*. ACM. <https://doi.org/10.1145/3593013.3594087>
- [45] Markus Langer, Cornelius J. König, Caroline Back, and Victoria Hemsing. 2022. Trust in Artificial Intelligence: Comparing Trust Processes Between Human and Automated Trustees in Light of Unfair Bias. *Journal of Business and Psychology* 38, 3 (June 2022), 493–508. <https://doi.org/10.1007/s10869-022-09829-9>
- [46] Markus Langer, Cornelius J. König, and Vivien Busch. 2020. Changing the means of managerial work: effects of automated decision support systems on personnel selection tasks. *Journal of Business and Psychology* 36, 5 (Sept. 2020), 751–769. <https://doi.org/10.1007/s10869-020-09711-6>
- [47] Markus Langer and Richard N. Landers. 2021. The future of artificial intelligence at work: A review on effects of decision automation and augmentation on workers targeted by algorithms and third-party observers. *Computers in Human Behavior* 123 (Oct. 2021), 106878. <https://doi.org/10.1016/j.chb.2021.106878>
- [48] Markus Langer, Daniel Oster, Timo Speith, Holger Hermanns, Lena Kästner, Eva Schmidt, Andreas Sasing, and Kevin Baum. 2021. What do we want from Explainable Artificial Intelligence (XAI)? – A stakeholder perspective on XAI and a conceptual model guiding interdisciplinary XAI research. *Artificial Intelligence* 296 (July 2021), 103473. <https://doi.org/10.1016/j.artint.2021.103473>
- [49] Min Kyung Lee. 2018. Understanding perception of algorithmic decisions: Fairness, trust, and emotion in response to algorithmic management. *Big Data & Society* 5, 1 (Jan. 2018), 205395171875668. <https://doi.org/10.1177/2053951718756684>
- [50] Gabriel Lima, Nina Grgić-Hlača, and Meeyoung Cha. 2021. Human Perceptions on Moral Responsibility of AI: A Case Study in AI-Assisted Bail Decision-Making. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems (CHI '21)*. Association for Computing Machinery, New York, NY, USA, 1–17. <https://doi.org/10.1145/3411764.3445260>
- [51] David Lindlbauer, Anna Maria Feit, and Otmar Hilliges. 2019. Context-Aware Online Adaptation of Mixed Reality Interfaces. In *Proceedings of the 32nd Annual ACM Symposium on User Interface Software and Technology (New Orleans, LA, USA) (UIST '19)*. Association for Computing Machinery, New York, NY, USA, 147–160. <https://doi.org/10.1145/3332165.3347945>
- [52] Han Liu, Vivian Lai, and Chenhao Tan. 2021. Understanding the Effect of Out-of-distribution Examples and Interactive Explanations on Human-AI Decision Making. *Proceedings of the ACM on Human-Computer Interaction* 5, CSCW2 (Oct. 2021), 1–45. <https://doi.org/10.1145/3479552>
- [53] Zhuoran Lu, Dakuo Wang, and Ming Yin. 2024. Does More Advice Help? The Effects of Second Opinions in AI-Assisted Decision Making. *Proc. ACM Hum.-Comput. Interact.* 8, CSCW1 (April 2024), 217:1–217:31. <https://doi.org/10.1145/3653708>
- [54] Shuai Ma, Xinru Wang, Ying Lei, Chuhan Shi, Ming Yin, and Xiaojuan Ma. 2024. “Are You Really Sure?” Understanding the Effects of Human Self-Confidence Calibration in AI-Assisted Decision Making. In *Proceedings of the CHI Conference on Human Factors in Computing Systems (CHI '24)*. Association for Computing Machinery, New York, NY, USA, 1–20. <https://doi.org/10.1145/3613904.3642671>
- [55] Frank Marcinkowski, Kimon Kieslich, Christopher Starke, and Marco Lünich. 2020. Implications of AI (un-)fairness in higher education admissions: the effects of perceived AI (un-)fairness on exit, voice and organizational reputation. In *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency (FAT* '20)*. ACM. <https://doi.org/10.1145/3351095.3372867>
- [56] Christian Meske and Erdi Ünal. 2024. Investigating the Impact of Control in AI-Assisted Decision-Making - An Experimental Study. In *Proceedings of Mensch und Computer 2024 (MuC '24)*. Association for Computing Machinery, New York, NY, USA, 419–423. <https://doi.org/10.1145/3670653.3677476>
- [57] Frederick P. Morgeson and Michael A. Campion. 2003. Work Design. , 423–452 pages. <https://doi.org/10.1002/0471264385.wei1217>
- [58] Frederick P. Morgeson and Stephen E. Humphrey. 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 (2006), 1321–1339. <https://doi.org/10.1037/0021-9010.91.6.1321>
- [59] Stephan J. Motowidlo, John S. Packard, and Michael R. Manning. 1986. Occupational stress: Its causes and consequences for job performance. *Journal of Applied Psychology* 71, 4 (1986), 618–629. <https://doi.org/10.1037/0021-9010.71.4.618>
- [60] Mareike Möhlmann, Lior Zalmanson, Ola Henfridsson, and Robert Wayne Gregory. 2021. Algorithmic Management of Work on Online Labor Platforms: When Matching Meets Control. *MIS Quarterly* 45, 4 (Oct. 2021), 1999–2022. <https://doi.org/10.25300/misq/2021/15333>
- [61] Kevin Nolan, Nathan Carter, and Dev Dalal. 2016. Threat of Technological Unemployment: Are Hiring Managers Discounted for Using Standardized Employee Selection Practices? *Personnel Assessment and Decisions* 2, 1 (July 2016). <https://doi.org/10.25035/pad.2016.004>
- [62] Raja Parasuraman and Dietrich H. Manzey. 2010. Complacency and Bias in Human Use of Automation: An Attentional Integration. *Human Factors: The Journal of the Human Factors and Ergonomics Society* 52, 3 (June 2010), 381–410. <https://doi.org/10.1177/0018720810376055>
- [63] Saumya Pareek, Eduardo Velloso, and Jorge Goncalves. 2024. Trust Development and Repair in AI-Assisted Decision-Making during Complementary Expertise. In *Proceedings of the 2024 ACM Conference on Fairness, Accountability, and Transparency (FAccT '24)*. Association for Computing Machinery, New York, NY, USA, 546–561. <https://doi.org/10.1145/3630106.3658924>
- [64] Xavier Parent-Rochelleau and Sharon K. Parker. 2022. Algorithms as work designers: How algorithmic management influences the design of jobs. *Human Resource Management Review* 32, 3 (Sept. 2022), 100838. <https://doi.org/10.1016/j.hrmr.2021.100838>
- [65] Gordon B. Parker and Matthew P. Hyett. 2011. Measurement of Well-Being in the Workplace: The Development of the Work Well-Being Questionnaire. *The Journal of Nervous and Mental Disease* 199, 6 (June 2011), 394. <https://doi.org/10.1097/NMD.0b013e31821cd3b9>
- [66] Sharon K. Parker. 2014. Beyond Motivation: Job and Work Design for Development, Health, Ambidexterity, and More. *Annual Review of Psychology* 65, 1 (Jan. 2014), 661–691. <https://doi.org/10.1146/annurev-psych-010213-115208>

- [67] Sharon K. Parker and Gudela Grote. 2020. Automation, Algorithms, and Beyond: Why Work Design Matters More Than Ever in a Digital World. *Applied Psychology* 71, 4 (Feb. 2020), 1171–1204. <https://doi.org/10.1111/apps.12241>
- [68] Sharon K. Parker and Caroline Knight. 2023. The <scp>SMART</scp> model of work design: A higher order structure to help see the wood from the trees. *Human Resource Management* 63, 2 (Nov. 2023), 265–291. <https://doi.org/10.1002/hrm.22200>
- [69] Mario Passalacqua, Robert Pellerin, Esma Yahia, Florian Magnani, Frédéric Rosin, Laurent Joblot, and Pierre-Majorique Léger. 2024. Practice With Less AI Makes Perfect: Partially Automated AI During Training Leads to Better Worker Motivation, Engagement, and Skill Acquisition. *International Journal of Human–Computer Interaction* 0, 0 (March 2024), 1–21. <https://doi.org/10.1080/10447318.2024.2319914> Publisher: Taylor & Francis _eprint: <https://doi.org/10.1080/10447318.2024.2319914>.
- [70] Forough Poursabzi-Sangdeh, Daniel G. Goldstein, Jake M. Hofman, Jennifer Wortman Vaughan, and Hanna Wallach. 2021. Manipulating and Measuring Model Interpretability. <http://arxiv.org/abs/1802.07810> arXiv:1802.07810 [cs].
- [71] Charvi Rastogi, Yunfeng Zhang, Dennis Wei, Kush R. Varshney, Amit Dhurandhar, and Richard Tomsett. 2022. Deciding Fast and Slow: The Role of Cognitive Biases in AI-assisted Decision-making. *Proc. ACM Hum.-Comput. Interact.* 6, CSCW1 (April 2022), 83:1–83:22. <https://doi.org/10.1145/3512930>
- [72] Vincent Robbmond, Oana Inel, and Ujwal Gadiraju. 2022. Understanding the Role of Explanation Modality in AI-assisted Decision-making. In *Proceedings of the 30th ACM Conference on User Modeling, Adaptation and Personalization (UMAP '22)*. Association for Computing Machinery, New York, NY, USA, 223–233. <https://doi.org/10.1145/3503252.3531311>
- [73] Richard M. Ryan and Edward L. Deci. 2000. Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist* 55, 1 (2000), 68–78. <https://doi.org/10.1037//0003-066x.55.1.68>
- [74] Richard M. Ryan and Edward L. Deci. 2017. *Self-Determination Theory: Basic Psychological Needs in Motivation, Development, and Wellness*. Guilford Press. <https://doi.org/10.1521/978.14625/28806>
- [75] Jakob Schoeffer, Maria De-Arteaga, and Niklas Kühl. 2024. Explanations, Fairness, and Appropriate Reliance in Human-AI Decision-Making. In *Proceedings of the CHI Conference on Human Factors in Computing Systems (CHI '24)*. ACM, 1–18. <https://doi.org/10.1145/3613904.3642621>
- [76] Shital Shah, Debadepta Dey, Chris Lovett, and Ashish Kapoor. 2018. Airsim: High-fidelity visual and physical simulation for autonomous vehicles. In *Field and Service Robotics: Results of the 11th International Conference*. Springer, 621–635.
- [77] Gretchen M. Spreitzer. 1995. An empirical test of a comprehensive model of intrapersonal empowerment in the workplace. *American Journal of Community Psychology* 23, 5 (Oct. 1995), 601–629. <https://doi.org/10.1007/BF02506984>
- [78] Peter J. Stanley, Nicola S. Schutte, and Wendy J. Phillips. 2021. A meta-analytic investigation of the relationship between basic psychological need satisfaction and affect. *Journal of Positive School Psychology* 5, 1 (April 2021), 1–16. <https://doi.org/10.47602/jpsp.v5i1.210>
- [79] Sarah Sterz, Kevin Baum, Sebastian Biewer, Holger Hermanns, Anne Lauber-Rönsberg, Philip Meinel, and Markus Langer. 2024. On the Quest for Effectiveness in Human Oversight: Interdisciplinary Perspectives. In *The 2024 ACM Conference on Fairness, Accountability, and Transparency (FAccT '24, Vol. 15)*. ACM, 2495–2507. <https://doi.org/10.1145/3630106.3659051>
- [80] Mark Steyvers and Aakriti Kumar. 2023. Three Challenges for AI-Assisted Decision-Making. *Perspectives on Psychological Science* (2023), 17456916231181102. <https://doi.org/10.1177/17456916231181102> arXiv:<https://doi.org/10.1177/17456916231181102> PMID: 37439761.
- [81] Barbara Stiglbauer and Carrie Kovacs. 2018. The more, the better? Curvilinear effects of job autonomy on well-being from vitamin model and PE-fit theory perspectives. *Journal of Occupational Health Psychology* 23, 4 (Oct. 2018), 520–536. <https://doi.org/10.1037/ocp0000107>
- [82] Franz Strich, Anne-Sophie Mayer, and Marina Fiedler. 2021. What Do I Do in a World of Artificial Intelligence? Investigating the Impact of Substitutive Decision-Making AI Systems on Employees' Professional Role Identity. *Journal of the Association for Information Systems* 22, 2 (2021), 304–324. <https://doi.org/10.17705/1jais.00663>
- [83] Siddharth Swaroop, Zana Bućinca, Krzysztof Z. Gajos, and Finale Doshi-Velez. 2024. Accuracy-Time Tradeoffs in AI-Assisted Decision Making under Time Pressure. In *Proceedings of the 29th International Conference on Intelligent User Interfaces (IUI '24)*. Association for Computing Machinery, New York, NY, USA, 138–154. <https://doi.org/10.1145/3640543.3645206>
- [84] Minmin Tang, Dahua Wang, and Alain Guerrien. 2019. A systematic review and meta-analysis on basic psychological need satisfaction, motivation, and well-being in later life: Contributions of self-determination theory. *PsyCh Journal* 9, 1 (June 2019), 5–33. <https://doi.org/10.1002/pchj.293>
- [85] Louis Tay, Cassandra Batz-Barbarich, Liu-Qin Yang, and Christopher W Wiese. 2023. Well-Being: The Ultimate Criterion for Organizational Sciences. (Aug. 2023). <https://doi.org/10.31234/osf.io/4v2y5>
- [86] Anna-Sophie Ulfert, Conny H. Antoni, and Thomas Ellwart. 2022. The role of agent autonomy in using decision support systems at work. *Computers in Human Behavior* 126 (2022), 106987. <https://doi.org/10.1016/j.chb.2021.106987>
- [87] Anja Van den Broeck, D. Lance Ferris, Chu-Hsiang Chang, and Christopher C. Rosen. 2016. A Review of Self-Determination Theory's Basic Psychological Needs at Work. *Journal of Management* 42, 5 (March 2016), 1195–1229. <https://doi.org/10.1177/0149206316632058>
- [88] Anja Van den Broeck, Maarten Vansteenkiste, Hans De Witte, Bart Soenens, and Willy Lens. 2010. Capturing autonomy, competence, and relatedness at work: Construction and initial validation of the Work-related Basic Need Satisfaction scale. *Journal of Occupational and Organizational Psychology* 83, 4 (Dec. 2010), 981–1002. <https://doi.org/10.1348/096317909x481382>
- [89] Rolf Van Dick, Christiane Schnitger, Carla Schwartzmann-Buchelt, and Ulrich Wagner. 2001. Der Job Diagnostic Survey im Bildungsbereich: Eine Überprüfung der Gültigkeit des Job Characteristics Model bei Lehrerinnen und Lehrern, Hochschulangehörigen und Erzieherinnen mit berufsspezifischen Weiterentwicklungen des JDS. *Zeitschrift für Arbeits- und Organisationspsychologie A&O* 45, 2 (April 2001), 74–92. <https://doi.org/10.1026//0932-4089.45.2.74>

- [90] Regina Vollmeyer and Falko Rheinberg. 2000. Does motivation affect performance via persistence? *Learning and Instruction* 10, 4 (Aug. 2000), 293–309. [https://doi.org/10.1016/s0959-4752\(99\)00031-6](https://doi.org/10.1016/s0959-4752(99)00031-6)
- [91] Xinru Wang and Ming Yin. 2021. Are Explanations Helpful? A Comparative Study of the Effects of Explanations in AI-Assisted Decision-Making. In *26th International Conference on Intelligent User Interfaces*. ACM, College Station TX USA, 318–328. <https://doi.org/10.1145/3397481.3450650>
- [92] Xinru Wang and Ming Yin. 2023. Watch Out for Updates: Understanding the Effects of Model Explanation Updates in AI-Assisted Decision Making. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems (CHI '23)*. Association for Computing Machinery, New York, NY, USA, 1–19. <https://doi.org/10.1145/3544548.3581366>
- [93] Elvin Yao and Jason T. Siegel. 2021. The influence of perceptions of intentionality and controllability on perceived responsibility: Applying attribution theory to people's responses to social transgression in the COVID-19 pandemic. *Motivation Science* 7, 2 (June 2021), 199–206. <https://doi.org/10.1037/mot0000220> Publisher: Educational Publishing Foundation.
- [94] John Zerilli, Alistair Knott, James Maclaurin, and Colin Gavaghan. 2019. Algorithmic Decision-Making and the Control Problem. *Minds and Machines* 29, 4 (Dec. 2019), 555–578. <https://doi.org/10.1007/s11023-019-09513-7>
- [95] Yunfeng Zhang, Q. Vera Liao, and Rachel K. E. Bellamy. 2020. Effect of confidence and explanation on accuracy and trust calibration in AI-assisted decision making. In *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency*. ACM, Barcelona Spain, 295–305. <https://doi.org/10.1145/3351095.3372852>

A Additional Questionnaire Items

Table 5. Exploratory questionnaire items and measured concepts

Concept	Question(s)	Ref
Personal Responsibility	'I was responsible for the crash of the drone.'	[37]
Causality	'I caused the drone to crash.'	[26]
	'I could have avoided crashing the drone.'	
	'I could have prevented the drone from crashing.'	
Knowledge	'I could have foreseen the crash of the drone.'	[26]
	'I understand why the drone crashed.'	
Intention	'I had the intention of crashing the drone.'	[26]
Control	'I felt like I was in control over the crash of the drone.'	[61]
Trust	'I can trust the system.'	[43]
	'I can rely on the available options.'	
Disengagement	'I feel that I have become disconnected from my work task. '	[19]
Exhaustion	'I feel tired performing the task. '	[19]
Stress	'I am currently stressed by the task.'	[59]
Blame	'To what extent do you think you are to blame for the drone crash?	[2]
	'To what extent do you think the system is to blame for the drone crash?'	
	'To what extent do you think you should be morally condemned for the crash of the drone?'	
Punishment	'To what extent should you be punished for the crash of the drone?'	[93]
	'To what extent should the provider of the system be punished for the crash of the drone?'	
Performance	'I was successful in accomplishing the task.'	[35]
Physical Demand	'The task was physically demanding.'	
Effort	'I had to work hard to accomplish my level of performance.'	
Frustration	'I was irritated, stressed, and annoyed during the task.'	