Computers in Human Behavior 144 (2023) 107714

Contents lists available at ScienceDirect

Computers in Human Behavior

journal homepage: www.elsevier.com/locate/comphumbeh

Full length article

Decision control and explanations in human-AI collaboration: Improving user perceptions and compliance

Monika Westphal [a,*,1], Michael Vössing [b,1], Gerhard Satzger [b], Galit B. Yom-Tov [c], Anat Rafaeli [c]

[a] University of Cologne, Faculty of Management, Economics and Social Sciences, Albertus-Magnus-Platz, 50923 Cologne, Germany
[b] Karlsruhe Institute of Technology, Institute of Information Systems and Marketing, Kaiserstraße 89, 76133 Karlsruhe, Germany
[c] Technion – Israel Institute of Technology, Faculty of Data and Decision Sciences, 3200003 Haifa, Israel

ARTICLE INFO

Keywords:
Human-AI collaboration
Decision control
Explanations
User trust
User compliance
Task complexity

ABSTRACT

Human-AI collaboration has become common, integrating highly complex AI systems into the workplace. Still, it is often ineffective; impaired perceptions – such as low trust or limited understanding – reduce compliance with recommendations provided by the AI system. Drawing from cognitive load theory, we examine two techniques of human-AI collaboration as potential remedies. In three experimental studies, we grant users decision control by empowering them to adjust the system's recommendations, and we offer explanations for the system's reasoning. We find decision control positively affects user perceptions of trust and understanding, and improves user compliance with system recommendations. Next, we isolate different effects of providing explanations that may help explain inconsistent findings in recent literature: while explanations help reenact the system's reasoning, they also increase task complexity. Further, the effectiveness of providing an explanation depends on the specific user's cognitive ability to handle complex tasks. In summary, our study shows that users benefit from enhanced decision control, while explanations – unless appropriately designed for the specific user – may even harm user perceptions and compliance. This work bears both theoretical and practical implications for the management of human-AI collaboration.

1. Introduction

The capabilities of artificial intelligence (AI) are continuously increasing. Humans can delegate a growing number of decision-making tasks to automated systems (Nunes & Jannach, 2017). AI can be used to reduce business costs, speed up existing processes, and improve the quality and effectiveness of managerial decisions (Ebel, Söllner, Leimeister, Crowston, & de Vreede, 2021). In many industries, AI now completes tasks that previously required the expertise of humans (Grace, Salvatier, Dafoe, Zhang, & Evans, 2018). In the coming years, the ratio of tasks performed by machines will increase dramatically (Cann, 2018). Hence, we expect to see an increase in human-AI collaboration through the adoption of systems that provide users with intelligent assistance to improve decision-making quality (Tariq & Rafi, 2012). This raises questions about people's willingness to trust and rely on these systems and, specifically, on their recommendations (Nunes & Jannach, 2017). The reasoning of state-of-the-art black-box models is inherently opaque and, thus, more challenging to comprehend than that of models used in traditional decision support systems – even for domain experts (Lundberg & Lee, 2017). This makes users reluctant to accept the recommendations of those models – a situation compounded by a general mistrust towards algorithmic decision-making (Logg, Minson, & Moore, 2019). Researchers increasingly argue that understanding "why a model makes a certain prediction can be as crucial as the prediction's accuracy itself" (Lundberg & Lee, 2017, p. 4766). Hence, models should be both accurate and comprehensible, enabling users to understand the reasoning of how they consider different input features to make a recommendation (Lundberg et al., 2020).

Significant effort has been invested to develop techniques that improve user outcomes in human-AI collaboration and, more specifically, in algorithmic decision-making. In this work, we study how users perceive and to what extent they comply with AI system recommendations. User perceptions and behaviors have been identified by many studies as important factors in human–computer interaction (HCI), precisely because they relate to user performance (Liang, Lee, & Jang, 2013; Tzafilkou & Protogeros, 2017). In this work, we use the term "user perceptions" to refer to various perceptual processes that users experience when collaborating. HCI research reports on a series of behavioral

* Corresponding author.
E-mail addresses: monika.westphal@uni-koeln.de (M. Westphal), michael.voessing@kit.edu (M. Vössing), gerhard.satzger@kit.edu (G. Satzger), gality@technion.ac.il (G.B. Yom-Tov), anatr@technion.ac.il (A. Rafaeli).
[1] These authors contributed equally.

https://doi.org/10.1016/j.chb.2023.107714
Received 26 January 2022; Received in revised form 24 January 2023; Accepted 14 February 2023
Available online 16 February 2023
0747-5632/

theories that shed light on both user perceptions (e.g., trust, perceived usefulness) and user behaviors (e.g., acceptance, use, adoption). One established theory is the Technology Acceptance Model (Venkatesh, Morris, Davis, & Davis, 2003). Recent empirical literature, investigating users' reactions to complex black-box models, has looked at few user perceptions so far, often in an isolated manner, (e.g., Bansal et al., 2021; Gutzwiller & Reeder, 2021; Haesevoets, De Cremer, Dierckx, & Van Hiel, 2021; Schlicker et al., 2021); among them are trust, perceived ease of use, and usefulness (Bansal et al., 2021; Shin, 2020; Tzafilkou & Protogeros, 2017). The most prevalent representative of user behavior in the recent empirical literature is user acceptance, (e.g., Haesevoets et al., 2021; Shin, 2021; Tzafilkou & Protogeros, 2017). It is important to design and implement new approaches in order to capture and analyze both user perceptions and behaviors in HCI environments (Tzafilkou & Protogeros, 2017). Following this theme, we adapt two existing techniques to improve user perceptions and behaviors: increasing decision control and providing explanations.

Decision control is the degree of power that users may exert to adjust a system recommendation. This flexibility is known to improve user perceptions and behaviors (Burton, Stein, & Jensen, 2020; Dietvorst, Simmons, & Massey, 2018). Cognitive load theory (cf., Sweller, 1994) suggests that task outcomes improve as a response to increased control – as long as the cognitive load does not increase significantly (Benke, Gnewuch, & Maedche, 2022; Van Merrienboer, Schuurman, De Croock, & Paas, 2002). Providing explanations of the system's reasoning is another technique widely considered to be a viable means for improving user perceptions and behaviors (Bansal et al., 2021; Gregor & Benbasat, 1999; Miller, 2019). However, empirical evidence increasingly suggests that providing explanations to users often does not yield the desired outcomes and may even induce negative side effects (Bansal et al., 2021; Kizilcec, 2016; Zhang, Vera Liao, & Bellamy, 2020). For example, providing too much or unclear information may erode user trust (Ahn, Almaatouq, Gulabani, & Hosanagar, 2021; Kizilcec, 2016), and providing overly complex explanations may decrease user engagement and induce longer response times (Giboney, Brown, Lowry, & Nunamaker, 2015). These inconclusive findings likely occur because of influencing factors that have not yet been considered. Additionally, there is "no clear consensus on what constitutes a good explanation" (Nunes & Jannach, 2017; Nunes, Miles, Luck, & De Lucena, 2012, p. 395). Hence, a deeper understanding is needed of when and why explanations would induce certain user perceptions and behaviors – and when they would not. We address this call and draw on cognitive load theory to explain *why* user perceptions and behavior may be hampered as a response to a provided explanation (Berthold, Röder, Knörzer, Kessler, & Renkl, 2011; Sweller, 1994): providing explanations significantly increases perceived task complexity. Further, we examine whether users' cognitive ability can help compensate for this negative effect and may even improve user outcomes.

In this work, we run three experimental studies to test for the effect of decision control and providing an explanation (i.e., explanation presence) on a variety of user outcomes. Recent literature has focused on one or two specific user outcomes (e.g., Bansal et al., 2021; Gutzwiller & Reeder, 2021; Haesevoets et al., 2021), among them being trust, perceived ease of use, and usefulness. We follow the call for more comprehensible black-box models (Lundberg et al., 2020) and also test user perceptions of understanding. Besides user perceptions, we focus on user compliance to show that decision control affects another relevant user behavior apart from acceptance. So far, little research exists on computer-induced user compliance, focusing on user requests (e.g., Adam, Wessel, & Benlian, 2021; Cialdini & Goldstein, 2004; Liang et al., 2013). Our study is one of the first to investigate user compliance with AI system recommendations. We examine both intended and actual compliance and use validated measures for all predicted outcomes.

In sum, this work contributes to empirical HCI literature integrating cognitive load theory (cf., Hollender, Hofmann, Deneke, & Schmitz, 2010; Kirschner, Ayres, & Chandler, 2011; Van Gog & Ayres, 2009). First, we investigate two techniques commonly used in human-AI collaboration to improve user outcomes (i.e., decision control and explanation presence). Second, we consider a variety of user outcomes (i.e., perceptions of trust and understanding, as well as intended and actual compliance). Third, we examine both a mechanism (i.e., perceived task complexity) and a boundary condition (i.e., cognitive ability) to the effect of explanation presence on user outcomes. This work has direct implications for managing human-AI collaborations at the workplace: While users should generally be provided with enhanced decision control, organizations should consider the specific user's cognitive ability when designing explanations.

2. Theory and hypotheses

2.1. Automation and human-AI collaboration

Automated systems that assist users in completing tasks are now common and will likely become even more so due to the increasing capabilities of the underlying technology (Limerick, Coyle, & Moore, 2014). Automation refers to the full or partial replacement of a function previously carried out by a human domain expert (Parasuraman, Sheridan, & Wickens, 2000). In many domains, automated systems already outperform these domain experts, particularly on repetitive tasks requiring high levels of cognitive capacity (Schlicker et al., 2021). Automated systems can assist users across a wide continuum: the degree of automation ranges from no automation (i.e., no AI system exists to offer any assistance; the human must make all decisions manually) to full automation (i.e., the AI system decides and acts autonomously without the human) (Parasuraman et al., 2000). The spectrum between these two extremes bears the potential for real collaboration between humans and AI. In human-AI collaboration, the AI system and the user interact, share their prior knowledge, and ideally decide unanimously. The process aims to combine the complementary capabilities of the AI system and the human (the user), to achieve common goals more effectively (Hemmer, Schemmer, Vössing, & Kühl, 2021; Limerick et al., 2014). An example would be optical character recognition systems that transmit instances to humans when the quality of fully automated transcription is not sufficient. Depending on the level of automation, the system support is made more or less explicit to the user (Limerick et al., 2014). By allowing users to delegate decisions to the AI system, they can be relieved of the burden of repetitive, monotonous, or mundane decision-making (Floridi et al., 2018; Schlicker et al., 2021). While various advantages of human-AI collaboration have been identified (Ebel et al., 2021; Zhou et al., 2021), there are a number of challenges. Two prominent examples are automation bias – the human tendency to favor recommendations from automated systems and to ignore even correct decisions made by a human (Goddard, Roudsari, & Wyatt, 2012; Skitka, Mosier, & Burdick, 2000) – and algorithm aversion – the human tendency to prefer a human decision, even if it is clear that the automated system predicts more accurately (Burton et al., 2020; Dietvorst, Simmons, & Massey, 2015; Dietvorst et al., 2018; Renier, Schmid Mast, & Bekbergenova, 2021). Algorithm aversion and other biases can be overcome by enhancing the user's decision control (by giving them the freedom of modifying the recommendation of the automated system) (Dietvorst et al., 2018), and by providing explanations (Parasuraman et al., 2000).

In this work, we investigate human-AI collaboration in the context of hotel revenue management. Accurate price predictions are essential for profitable revenue management. AI systems can assist users (i.e., hotel managers) in determining suitable hotel room prices. This is a complex task, as the user would have to take various hotel characteristics into account as well as the competitors, the holiday season, etc. When studying human-AI collaboration (regardless of the context), biased data is an important issue to consider. We ensured that our AI system does not amplify racial bias (e.g., by removing any features from the data set like review texts). Further, we accounted for a possible selection bias by including all hotels provided, and reduced the risk of algorithmic bias by using 5-fold cross-validation when training the AI system. It consistently yielded high performance across different training folds and random initialization seeds.

Next, we will outline our rationale to analyze the effect of two techniques of human-AI collaboration (i.e., decision control and explanation presence) on the outcomes of human-AI collaboration (i.e., user perceptions and compliance). Following cognitive load theory (cf., Sweller, 1994), we first expect enhanced decision control to be an effective means for improving these user outcomes. Second, we propose that the effectiveness of providing explanations depends on multiple factors: perceived task complexity will explain the effect, while the specific user's cognitive ability will mitigate the effect. Each of the following subsections corresponds to one set of our hypotheses.

2.2. Effect of decision control on user perceptions and compliance

Decision control (or locus of control) generally refers to the extent to which people can modify the decisions or recommendations of a third party (Lee, Jain, Cha, Ojha, & Kusbit, 2019; Lind, Lissak, & Conlon, 1983; Thibaut & Walker, 1978). Similarly, allowing users to – even slightly – adjust the recommendations provided by an AI system (from now on: system) means delegating control over the decision to the user (Dietvorst et al., 2018). Researchers have long recognized control as a "key factor in how people perceive interactions with technology" (Limerick et al., 2014, p. 1). Decision control is a technique in human-AI collaboration (Dietvorst et al., 2018; Lee et al., 2019) that has the potential to improve user outcomes. Users generally desire to control the technologies they use (Legaspi & Toyoizumi, 2019) and expect systems to respond to their actions (Limerick et al., 2014). Accordingly, it is crucial to build systems that support the user's internal locus of control (Shneiderman & Plaisant, 2006).

Cognitive load theory considers people's limited cognitive processing capacity when learning and performing tasks (Paas, Tuovinen, Tabbers, & Van Gerven, 2003). Research shows that enhancing control can be accompanied by increased cognitive load (i.e., the mental effort people experience when learning or performing a task, see Paas and Van Merriënboer (1994)); for example, if too many response options are available (Kirschner et al., 2011; Van der Land, Schouten, Feldberg, van Den Hooff, & Huysman, 2013). Still, even if people experience increased cognitive load as a response to enhanced control, high performance can be achieved (Parasuraman, Sheridan, & Wickens, 2008; Van Merrienboer et al., 2002). Increased decision control has been shown to improve task performance by reducing extraneous and intrinsic load on the one hand and increasing germane load on the other hand (Van Merrienboer et al., 2002; Vandewaetere & Clarebout, 2013). These three types of cognitive load are all caused by the design and interaction with the task. Extrinsic load refers to the difficulty in learning with instructional material. Intrinsic load refers to the content's difficulty (Skulmowski & Xu, 2021). Hence, they both refer to unnecessary information processing (e.g., due to poor visualization), which inhibits learning. In contrast, germane load refers to the level of concentration and schema acquisition, thus encouraging learning (Skulmowski & Xu, 2021). Increased extraneous and/or intrinsic load causes a higher mental effort associated with the task, while increased germane load refocuses people's attention on the task (Van Merrienboer et al., 2002).

Enhanced decision control might increase people's task involvement so that they are more motivated to re-invest freed-up processing resources in learning (Van Merrienboer et al., 2002). This might explain why they do not perceive the mental effort associated with the task as higher and show not unchanged but improved performance. In contrast, when tasks are perceived as too difficult, people are less motivated to invest mental effort in order to complete the task (Vandewaetere & Clarebout, 2013). Research has revealed that increased decision control improves user outcomes: For example, Benke et al. (2022) show that, when interacting with chatbots, increased decision control improves perceptions of user trust and task performance, while cognitive load only increased slightly but not significantly. Further, Dietvorst et al. (2018) show that giving users the option to at least slightly adjust forecasting outcomes improves task performance. They also showed that users were more satisfied with the forecasting process, more likely to believe that the system was more accurate than the users themselves, and more likely to use it for subsequent forecasts. Based on all this, we propose:

H1a. High decision control improves user perceptions of (1) trust in, and (2) understanding of the system recommendation.

H1b. High decision control improves users' (1) intended, and (2) actual compliance with the system recommendation.

Fig. 1 presents the hypotheses for the proposed effect of decision control on user perceptions and compliance.

2.3. Effect of explanation presence on user perceptions and compliance

Human-AI collaboration requires that the AI is capable of communicating the knowledge it contains and of explaining its reasoning (Gregor & Benbasat, 1999). Automatically generated explanations have long been considered a fundamental mechanism to increase users' trust in provided recommendations (Nunes & Jannach, 2017). The presence or absence of explanations is an important design element of systems that are not fully automated (Papenmeier, Kern, Englebienne, & Seifert, 2022). Explanations are generally initiated by an information provider to clarify or justify the recommendation and convince the recipient to comply with it (Gregor & Benbasat, 1999). As an example, the inputs that were most decisive when determining the output can be explained (Nunes & Jannach, 2017; Shin & Park, 2019). Explanations are designed to increase transparency (Vössing, Potthoff, Kühl, & Satzger, 2019), and should increase trust in as well as induce user adoption of the provided recommendation (Berthold et al., 2011; Rader, Cotter, & Cho, 2018). Indeed, some empirical studies indicate that explanations improve user perceptions and behaviors (Gregor & Benbasat, 1999; Nunes & Jannach, 2017). In recent years, a substantial amount of work has been done on explanations in the context of advice-giving and expert systems (Nunes & Jannach, 2017). Due to the recent advances in the field of AI, providing explanations that are both correct and understandable has become more challenging; for example, when the recommendation of the system originates from a deep neural network (Nunes & Jannach, 2017; Shin & Park, 2019).

More recent studies have failed to find a positive effect of explanations on user perceptions and user behaviors (Bansal et al., 2021; Chandrasekaran, Prabhu, Yadav, Chattopadhyay, & Parikh, 2018; Hase & Bansal, 2020). Regardless of the system's complexity, the effectiveness of explanations also depends on the specific format or the way they are presented to the user (Gregor & Benbasat, 1999). Nunes and Jannach (2017) argue that even fine-grained details of the design can determine the effectiveness of explanations. Overall, more studies are needed to understand the advantages and disadvantages of different types of explanations (Nunes & Jannach, 2017). In this study, we focus on the effect of providing an explanation (i.e., explanation presence) on the user. Following the above-mentioned call, we include a robustness check to test whether the results hold if we change one element in the design of the provided explanation. We use a widely accepted type of explanation (i.e., Shapley values, see Lundberg et al. (2020)) and take the users' understanding of the explanation into account. Hence, overall, we expect to see a positive effect of providing an explanation on user perceptions and compliance. We propose the following:

H2a. Explanation presence improves user perceptions of (1) trust in, and (2) understanding of the system recommendation.

H2b. Explanation presence improves users' (1) intended, and (2) actual compliance with the system recommendation.

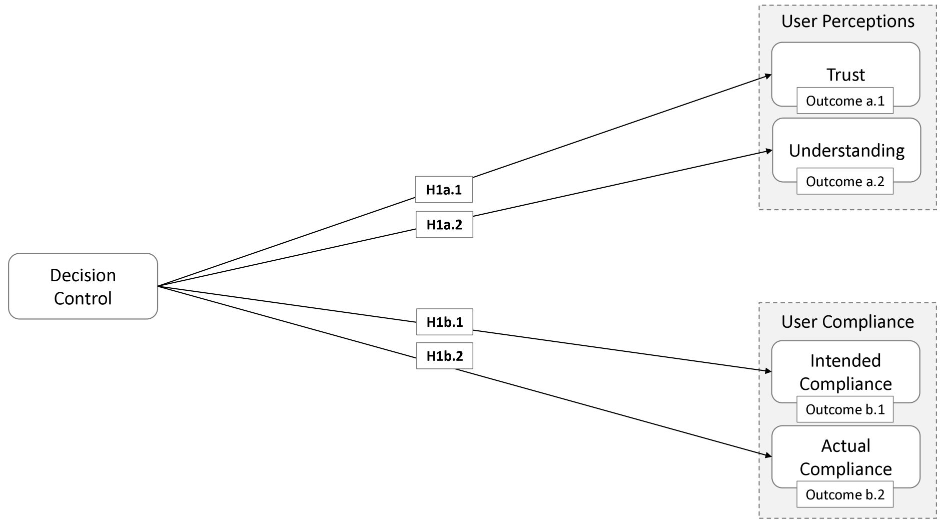


Fig. 1. Research model I: Proposed effects of decision control on user perceptions and compliance.

2.4. Perceived task complexity as a mediator to the effect of explanation presence on user outcomes

Explanations are generally used with the implicit assumption that they reduce the complexity of a task by communicating the reasoning of the system. However, research in the context of e-learning has shown that providing explanations increases cognitive load when they reveal complex content (Berthold et al., 2011). Cognitive load theory states that people's cognitive capacity might be overloaded if the task environment is too rich in information (i.e., including multiple cues about various aspects of the task) (Paas et al., 2003). As a result, information might not be processed effectively and understanding might not be reached (Schrader & Bastiaens, 2012). Berthold et al. (2011) show that providing explanations can lead to double-edged effects – improving some task outcomes while impairing some others (e.g., understanding and performance).

Tasks are complex if they place high cognitive demands on the user (Campbell, 1988). Task complexity can impair the quality of users' decisions by reducing the accuracy of their judgment (Chan, Song, & Yao, 2015). Social cognitive theory suggests that users prefer simple over complex tasks because simple tasks entail more certain outcomes, which decrease the risk of failure (Chan et al., 2015). Perceived task complexity has been identified as a moderator to the effect of users' motivation on system use (Chan et al., 2015). In particular, if a task is perceived as simple, the user is less motivated to use the system because the task seems solvable without it (Chan et al., 2015). In contrast, if a task is perceived as complex, the user is more motivated to use the system. However, it is still unclear how users would react to an explanation that is in itself complex. Recent research indicates that a better understanding of influencing factors is needed to justify design decisions for explanations (Nunes & Jannach, 2017). Following this call, we examine the role of perceived task complexity in the effect of providing an explanation on user outcomes. We argue that a provided explanation increases rather than decreases users' cognitive demands required to solve the task: the more complex the task, the more information the explanation presents (simple tasks should not require any explanation to begin with). This additional information should signal task complexity rather than help reduce it. In sum, we propose that perceived task complexity mediates the effect of providing an explanation on user perceptions and compliance:

H3a. Explanation presence increases perceived task complexity, which in turn impairs users' (1) trust in, and (2) understanding of the system recommendation.

H3b. Explanation presence increases perceived task complexity, which in turn impairs users' (1) intended, and (2) actual compliance with the system recommendation.

2.5. Cognitive ability as a moderator to the effect of explanation presence on user outcomes

To further explore how explanations affect users – as suggested by Nunes and Jannach (2017) – we also investigate cognitive ability as a potential boundary condition to the effect. Researchers have tried to understand the connection between cognitive information processing capabilities and organizational behavior (Woznyj, Banks, Dunn, Berka, & Woehr, 2020). Research has shown that users' characteristics can influence their reactions to a provided explanation (Gregor & Benbasat, 1999), including expertise (Mao, Sen, & Turner, 2018) and cognitive style (i.e., the way users think, perceive, and remember information) (Hsu, 1993).

Cognitive load theory has focused on users' expertise as a factor in reducing cognitive load – thus yielding positive task outcomes (Hollender et al., 2010; Van Gog & Ayres, 2009). We investigate human-AI collaboration in a context where users typically have not yet encountered any AI systems and hence have not gained any expertise. Given that users are no experts, we are interested in investigating user characteristics that are inherent to the individual, independent from the specific task, and yet potentially affected by a provided explanation. We focus on the specific user's cognitive ability in terms of handling complex tasks. Cognitive ability refers to the ability to reason and solve problems (Woznyj et al., 2020). Research suggests that cognitive ability is crucial when encountering a new task; over time, experience and expertise become more important than cognitive ability (Kanfer & Ackerman, 1989). Cognitive ability is known to predict job performance in various settings (Schmidt, 2002). Further, it has been claimed that people with high cognitive ability might be willing to adopt new technology, even when facing high task complexity (Mun, Jackson, Park, & Probst, 2006). Similarly, if the amount of information processing exceeds a certain limit, people's cognitive ability moderates performance (Norman & Bobrow, 1975). We expect that users with high cognitive ability will react more positively to a provided explanation embedded in a complex task. In other words, we propose that cognitive ability moderates the effect of providing an explanation on user perceptions and compliance:

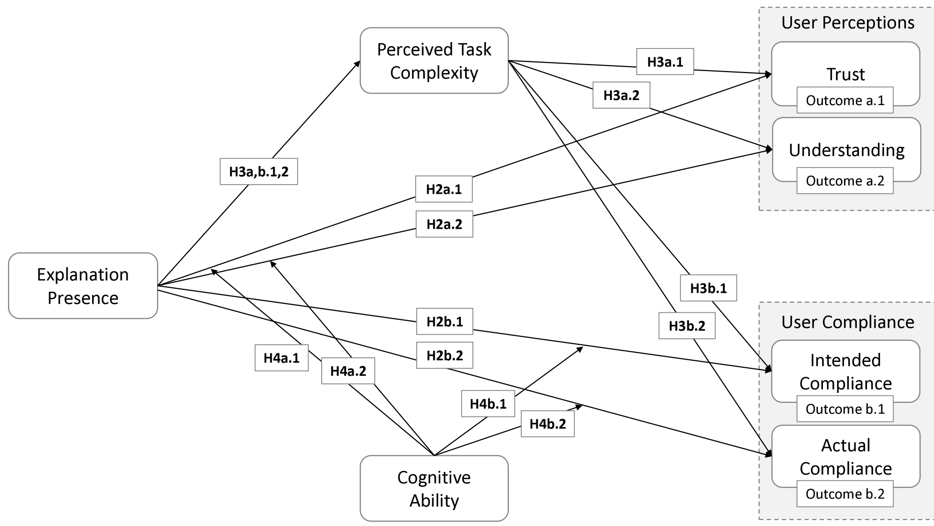


Fig. 2. Research model II: Proposed effects of explanation presence, perceived task complexity, and cognitive ability on user perceptions and compliance.

H4a. The higher the cognitive ability, the stronger explanation presence improves users' (1) trust in, and (2) understanding of the system recommendation.

H4b. The higher the cognitive ability, the stronger explanation presence improves users' (1) intended, and (2) actual compliance with the system recommendation.

Fig. 2 presents all hypotheses for the proposed effects of explanation presence, perceived task complexity and cognitive ability on user perceptions and compliance.

3. Overview of the studies

We conducted three online experimental studies to test our hypotheses. The Institutional Review Board of the Israel Institute of Technology approved the line of studies. Additionally, written informed consent was obtained from all participants prior to inclusion. All three studies examined human-AI collaboration in the context of hotel revenue management, and users (i.e., hotel managers) encounter an AI system when determining suitable hotel room prices. In Study 1, we tested whether decision control improves user outcomes (perceptions and compliance) – i.e., Hypotheses 1a and 1b. In Studies 2 and 3, we focused on the effectiveness of providing an explanation. In Study 2, we tested whether explanation presence improves user outcomes (Hypotheses 2a and 2b). Study 3 served two purposes: First, it provided a robustness check for the direct effect of explanation presence on user outcomes. Second, it analyzed whether these effects are mediated by perceived task complexity (Hypotheses 3a and 3b) and moderated by the specific user's cognitive ability (Hypotheses 4a and 4b). Next, we describe the chosen method for each study, present the study's results, and discuss them in light of our overarching research questions.

4. Study 1: Examining the effect of decision control on user perceptions and compliance

4.1. Method

Participants

113 Participants were recruited online via the Prolific platform for a 15-min study on estimating room prices for hotels and received $2 for their participation. Prolific is an online platform that provides researchers access to a diverse participant pool and acts as a trusted intermediary (Peer, Brandimarte, Samat, & Acquisti, 2017). All participants had at least some work experience in the hospitality and tourism sector. We calculated the required sample size using G-Power (Faul, Erdfelder, Lang, & Buchner, 2007) and assumed a moderate effect (.15). Accordingly, 107 participants were necessary to detect effects between two groups in a multiple linear regression (fixed model, R^2 deviation from zero), with a power of .95. As it is common for some participants to fail attention tests or drop out of the study, we recruited a slightly larger number of participants. Indeed, some participants dropped out, and some were excluded based on failed attention or manipulation checks, as well as for problematic values; see Table 1. Finally, 110 participants were included in the analysis ($M = 29$ years, $SD = 11.4$; 66% female).

Procedure

Participants were first introduced to the study context: we informed them that revenue managers of hotels often struggle to determine the prices they should charge for rooms based on the hotels' characteristics (e.g., star rating, location, distance to the city center, and customer reviews). Next, participants had to pass an attention check. Then, they were asked to estimate the room price for 6 hotels located in London (*Task I* – initial decision) based on the specific values of 13 characteristics for each hotel (see Fig. C.10). They were also informed that the average price for a hotel room in London was £180.

Next, we informed participants that a machine learning model was available to predict the room prices and provided them with the

Table 1
Data exclusion (Studies 1–3).

Study	Initial sample size	Dropped out (%)	Failed ATC (%)	Failed MC (%)	Problematic values[1]	Final sample size
1	113	0 (0%)	2 (2%)	–	1 (1%)	110
2	120	9 (7%)	0 (0%)	–	1 (1%)	110
3	315	12 (4%)	6 (2%)	27 (9%)	7 (2%)	263

Notes.[1] Task I or Task II; either missing, or negative, or >3.5SD.
ATC = Attention Check, MC = Manipulation Check.

model's recommendation (i.e., predicted room price) for each of the 6 hotels – in addition to the specific values of the 13 hotel characteristics and the average room price (see an example for one hotel in Fig. C.11). Then, participants were again asked to estimate the room price for the 6 hotels (*Task II*– updated decision). Their new price estimate could be either identical to their previous estimation or different from it. Participants were randomly assigned to one of two *decision control* conditions (i.e., *low* vs. *high*). In the low decision control condition, participants could choose between two values: their initial decision (determined in Task I) or the model's recommendation. In the high decision control condition, participants could adjust their initial decision to any value. To measure each participant's actual compliance, we took the change from the initial to the updated decision regarding each of the 6 hotels as well as the average room price into account (see below). Following the main task, participants rated their perceptions (i.e., trust and understanding) of and their intention to comply with the system recommendation on validated scales. Lastly, we asked participants to report their demographics.

Measures

User perceptions. We assessed user perceptions with two constructs: *trust* and *perceived understanding*. Both constructs were measured with three items each on a Likert scale ranging from '1–I totally disagree' to '7–I totally agree'. For a full list of the items and statements used, see Table 2.

User compliance. We assessed user compliance with two constructs: *intended compliance* and *actual compliance*. Users' *intended compliance* was assessed with the construct *intention to use*, measured with two items on a Likert scale ranging from '1–I totally disagree' to '7–I totally agree' (see Table 2). In our study, use [of the model] refers to compliance [with the model recommendation].

Users' actual compliance with the model recommendation was defined as the "relative change in the accuracy of the participants' room price estimations from Task I (i.e., initial decision) to Task II (i.e., updated decision)". We defined the accuracy of a participant's room price estimation as "its deviance from the model's recommendation" (separately for each of the 6 hotels). To account for the large difference in the room prices of the 6 hotels (i.e., between £144 and £316), we used the *Mean Relative Absolute Error (MRAE)* (Hyndman & Koehler, 2006). MRAE is a ratio comparing the residual errors of a participant's estimations (compared to the model's recommendations) with those produced by a benchmark method (i.e., predicting the average room price for all hotels, £180, and again comparing to the model's recommendations). The metric is defined as $MRAE = mean(|e_t/e_t^*|)$, where e_t are the residual errors of the participant's estimations and e_t^* the errors obtained with the benchmark method. Due to the high accuracy of the AI (i.e., £21.29 mean average error (MAE) on the test set), it would be in the participants' best interest to accept the model recommendations in most cases. However, we do not communicate the model's accuracy. The relative change in the accuracy of the participants' room price estimations (i.e., $C = (MRAE_{after} - MRAE_{before})/MRAE_{before}$) measures the participant's actual compliance with the model recommendation. For easier understanding, we reversed the scale so that '+1' indicates the highest level of compliance.

Control variables. We included gender, age, education, and work experience in hospitality and revenue management and with revenue management software as control variables in all analyses.

Manipulation check. To ensure our manipulation for high and low decision control worked as intended, we ran a manipulation check ($N = 60$). The manipulation check was not part of Study 1, but we present the results here to maintain a consistent structure of the paper. Five participants failed the attention check and were excluded from the analysis. We asked participants about the extent to which they agree with the following statements (on a scale from '1–not at all' to '5–totally'): "I could modify the task outcome without restraint", "I could choose from many options when determining the task outcome" and "I had much freedom in determining the task outcome". Participants in the high decision control conditions perceived their decision control as significantly higher ($M = 4.20$, $SD = 0.75$) than in the low decision control condition ($M = 3.64$, $SD = 1.09$), $p = 0.033$. Hence, we concluded that our manipulation for decision control was valid.

Statistical specification

We tested the effect of decision control on user perceptions (trust and understanding: H1a.1 and H1a.2) and compliance (intended and actual: H1b.1 and H1b.2). We ran a separate multiple regression analysis for each of the four predicted outcomes. In all analyses, we controlled for participants' demographics.

4.2. Results

Effect of decision control on user perceptions

High decision control improved participants' trust ($p = 0.004$) and understanding ($p = 0.004$), compared to low decision control; see Fig. 3. Table 3 details the regression results. Hypotheses 1a.1 and 1a.2 were supported.

Effect of decision control on user compliance

High decision control improved participants' intended ($p = 0.029$) and actual compliance ($p = 0.010$), compared to low decision control (see Fig. 3 and Table 3). Hypotheses 1b.1 and 1b.2 were supported.

We ran a post hoc mediation analysis (Hayes, 2014, model no. 4) and found that participants more strongly complied with the system recommendation because they intended to comply more with it ($p = 0.007$), $F(10, 99) = 2.065$, $R^2 = 0.173$, $p = 0.034$, while controlling for trust and understanding.

4.3. Discussion

Study 1 shows that high decision control improves user perceptions of trust and understanding, as well as intended and actual compliance. This finding is in line with Dietvorst et al. (2018) and Benke et al. (2022), who demonstrated that giving people control over the task outcome improves perceptions, as well as task performance. In Study 2, we will test if explanation presence – another technique commonly used in human-AI collaboration – also improves user perceptions of trust and understanding, as well as intended and actual compliance.

Table 2
Measures of user perceptions and intended compliance (Studies 1–3).

Construct	Item	Statement	Source
Trust	T1	I am confident in the model.	Jian, Bisantz, and Drury (2000)
	T2	The model is reliable.	
	T3	I can trust the model.	
Understanding	UN1	The model is easy to understand.	Obar and Oeldorf-Hirsch (2020)
	UN2	The model provides helpful information.	
	UN3	The model is clear.	
Intended Compliance	IC1	Assuming I had access to the model, I would intend to use it in the job.	Venkatesh et al. (2003)
	IC2	Given I had access to the model, I predict that I would use it in the job.	

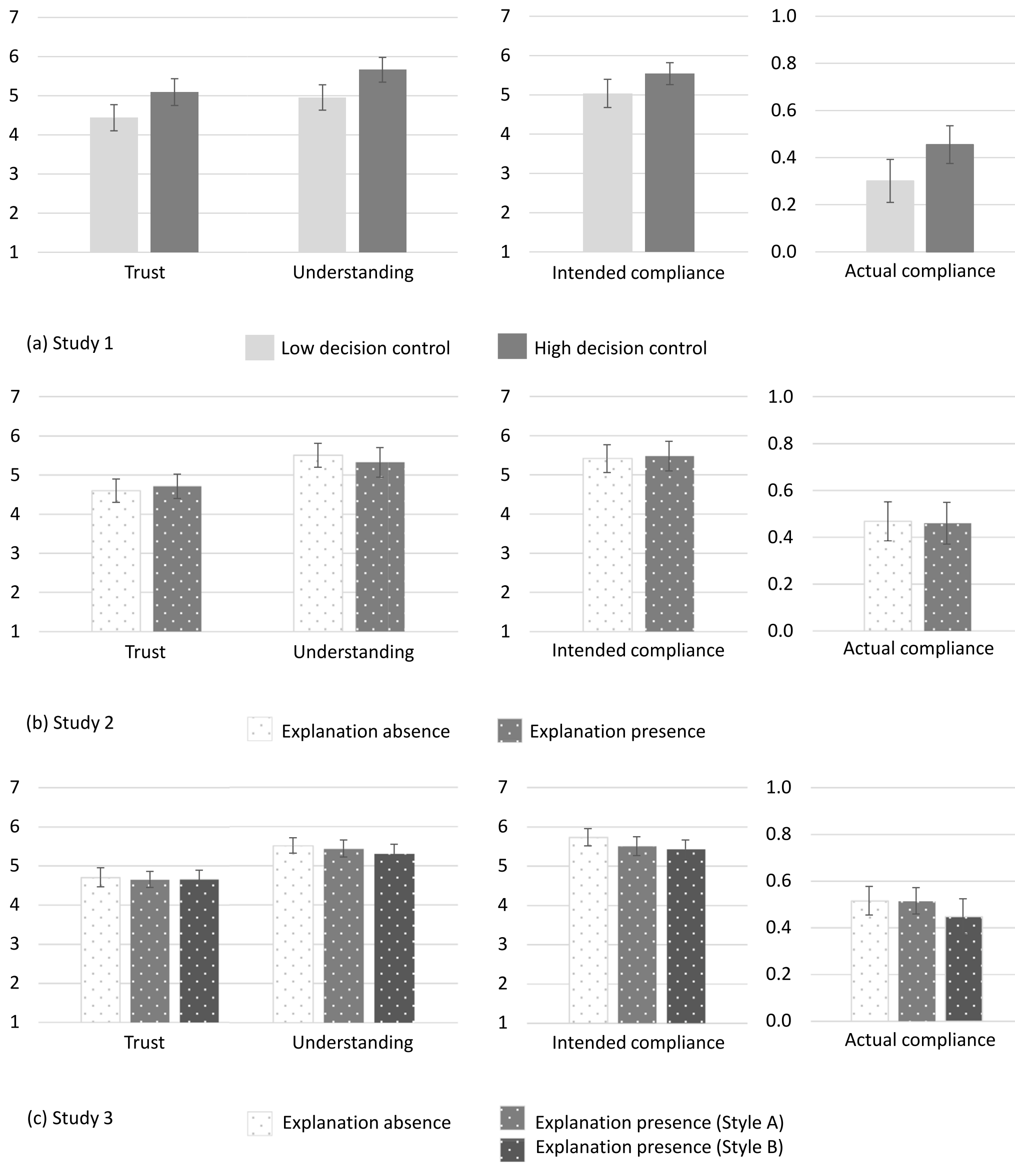


Fig. 3. Descriptives: Decision control, explanation presence, user perceptions and compliance (Study 1–3).

Table 3
Results: Effect of decision control on user perceptions and compliance (Study 1).

	Outcomes							
	User perceptions				User compliance			
	Trust		Understanding		Intended		Actual	
	coeff	*se*	*coeff*	*se*	*coeff*	*se*	*coeff*	*se*
Constant	5.25***	0.64	6.02***	0.65	5.96***	0.64	0.53**	0.17
Decision control	0.69**	0.24	0.71**	0.24	0.52*	0.24	0.17*	0.06
Gender	−0.28	0.25	−0.72**	0.26	−0.39	0.25	−0.09	0.07
Age	0.01	0.01	0.00	0.01	−0.01	0.01	0.00	0.00
Education	−0.20	0.13	−0.17	0.13	−0.10	0.13	−0.05	0.03
Experience hospitality mgmt.	0.03	0.07	0.04	0.07	−0.05	0.07	0.01	0.02
Experience revenue mgmt.	−0.02	0.10	0.01	0.10	0.15	0.10	−0.02	0.03
Experience revenue software	−0.42	0.38	−0.58	0.38	−0.34	0.38	−0.15	0.10
N	110		110		110		110	
R^2	0.113		0.166		0.108		0.107	
Adj. R^2	0.052		0.109		0.046		0.046	
Residual Std. Error	2.628		4.319		2.518		0.180	
F Statistic (df)	1.850† (7,102)		2.899** (7,102)		1.756 (7,102)		1.746 (7,102)	

Note. *** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$; † $p < 0.10$.

5. Study 2: Examining the effect of explanation presence on user perceptions and compliance

5.1. Method

Participants

120 Participants were recruited online via the Prolific platform for a 15-min study on estimating room prices for hotels and received $2 for their participation. All participants had at least some work experience in the hospitality and tourism sector. We calculated the required sample size using G-Power as in Study 1. 110 participants were included in the study (see Table 1) ($M = 29$ years, $SD = 9.0$; 65% female).

Procedure

The procedure of Study 2 was identical to Study 1, except for the following modifications: (1) In Task II, all participants were assigned to the high decision control condition, allowing them to adjust their initial decision to any value. (2) In Task II, participants were randomly assigned to an *explanation presence* condition (i.e., *presence* vs. *absence*). We presented explanations in a format based on the Shapley values (Lundberg et al., 2020; Lundberg & Lee, 2017) and communicated to users which inputs were most decisive when determining the output (Nunes & Jannach, 2017). Shapley-based explanations are commonly used in computer science to make decisions of black-box models more transparent for domain experts (Lundberg et al., 2020; Lundberg & Lee, 2017). However, they are hard for laypeople to understand and need to be adapted. Hence, in the current design, we simplified the Shapley-based explanations. The explanations communicated the marginal impact each hotel characteristic had on the room price recommended by the model. The explanation used in Study 2 (see Fig. C.12) presented the impacts of each hotel characteristic according to the arbitrary order of the hotel characteristics in Task I. Fig. C.11 shows the explanation-absence condition. (3) We ran a pretest ($N = 76$) to see whether participants actually understood the provided explanation and also asked for their perceived understanding of the explanation (see Section 5.2). We included both measures in Study 2.

Measures

This study included the same measures (for user perceptions and compliance) as Study 1, see also Table 2.

Control variables. In addition to the control variables in Study 1, this study included the following two: participants' perceived and actual understanding of the explanation. At the end of the study, all participants received an explanation they (might have) encountered in Task II. Regardless of whether they were assigned to the *explanation presence* or the *explanation absence* condition, we simply asked if they thought the explanation was easy to understand (on a scale from '1–not at all' to '5–totally'). In addition, we used the following three items to test participants' actual understanding of the explanation: "What is the predicted room price?" [£141/£160/£180], "The review location score [increased/decreased] the predicted room price, while star rating [increased/decreased] the predicted room price", and "The hotel has a [review location score of 9.2]/the review location score [increased/reduced the predicted price by £9.20]". For each of the three items, they also had the option to say "I don't know". We defined low explanation understanding as less than two (out of three) correct responses and high explanation understanding as three (out of three) correct responses.

Statistical specification

We tested for the effect of explanation presence on user perceptions (trust and understanding: H2a.1 and H2a.2) and compliance (intended and actual: H2b.1 and H2b.2), see Fig. 2. We again ran a separate multiple regression analysis for each of the four predicted outcomes. In all analyses, we controlled for participants' demographics. We also controlled for how well participants understood the explanation – or would have if they had received it – and participants' perceived understanding of the explanation.

5.2. Results

Pretest

We ran a pretest ($N = 77$) to examine the extent to which people understand the explanation we provide (cf., the measure of actual explanation understanding in Study 2). One participant was excluded due to missing values. We found that 66% (50/76) of the participants (objectively) fully understood the provided explanation. At the same time, only 15% (11/76) of the participants *felt* they fully understood it. On average, participants felt they moderately understood the provided explanation ($M = 3.22$, $SD = 1.17$). We also received qualitative feedback on the (seemingly too complex) design of the explanation, which we integrated into the design of the explanation for Study 2.

Effect of explanation presence on user perceptions

Despite the simplified design of the explanation, providing an explanation neither improved participants' trust ($p = 0.500$) nor their understanding ($p = 0.628$); see Fig. 3 and Table 4. Hypotheses 2a.1 and 2a.2 were not supported.

Effect of explanation presence on user compliance

Providing an explanation neither improved participants' intended ($p = 0.431$) nor their actual compliance ($p = 0.378$); see Fig. 3. Table 4 presents the regression results. Hypotheses 2b.1 and 2b.2 were not supported.

5.3. Discussion

Study 2 did not show evidence that providing an explanation improves user perceptions of trust and understanding, nor user intended or actual compliance. Following the pretest, we revised the – seemingly too complex – format of the explanation, resulting in a seven-point increase in people's understanding of the explanation. Still, only 73% fully understood it. We assume the explanation was still difficult to grasp for the participants. The following study will serve as a robustness check for Study 2 results, which are not in line with the – in practice, often intuitively expected – positive effect of explanation presence on user outcomes. Specifically, we will test if the results hold when we change one element in the (order the information is presented in the) explanation. Second, we will investigate if perceptions of increased task complexity explain the effect of explanation presence and whether the specific user's cognitive ability will moderate it.

6. Study 3: Examining perceived task complexity and cognitive ability for the effect of explanation presence on user outcomes

6.1. Method

Participants

315 Participants were recruited online via the Prolific platform for a 15-min study on estimating room prices for hotels and received $2 for their participation. All participants had at least some work experience in the hospitality and tourism sector. This study included additional predictors and measures. Hence, we assumed a small effect (.10). 262 participants were required to detect effects between three groups – explanation (Style A and Style B) presence and explanation absence – in a multiple linear regression (fixed model, R^2 deviation from zero). We additionally included a manipulation check for the explanation presence condition and recruited an even larger number of participants. Finally, 266 participants were included in the study (see Table 1) ($M = 31$ years, $SD = 10.5$; 64% female).

Procedure

The procedure of Study 3 was identical to Study 2, except for the following modifications: (1) At the beginning of the study, we additionally measured the specific participant's cognitive ability. (2) In Task II, participants were randomly assigned to one of three *explanation presence* conditions: *explanation (Style A)* (similar to the explanation provided in Study 2, with small graphical adjustments) vs. *explanation*

Table 4
Results: Effect of explanation presence on user perceptions and compliance (Study 2).

	Outcomes							
	User perceptions				User compliance			
	Trust		Understanding		Intended		Actual	
	coeff	*se*	*coeff*	*se*	*coeff*	*se*	*coeff*	*se*
Constant	3.25***	0.92	3.44***	0.97	3.41***	1.04	0.66**	0.24
Explanation presence	0.16	0.23	−0.12	0.25	0.21	0.27	0.06	0.06
Gender	−0.09	0.24	−0.06	0.25	−0.08	0.27	0.08	0.06
Age	−0.01	0.01	−0.02	0.02	−0.02	0.02	−0.02***	0.00
Education	0.11	0.11	0.10	0.12	0.08	0.13	0.01	0.03
Experience hospitality mgmt.	0.11	0.07	0.09	0.08	0.03	0.08	0.03	0.02
Experience revenue mgmt.	−0.08	0.11	−0.04	0.11	0.16	0.12	0.08**	0.03
Experience revenue software	0.03	0.41	0.08	0.43	0.09	0.46	−0.02	0.11
Act. understand. explanation	0.28	0.26	−0.05	0.27	−0.19	0.29	0.05	0.07
Perc. understand. explanation	0.24*	0.11	0.47***	0.12	0.47***	0.13	0.01	0.03
N	110		110		110		110	
R^2	0.082		0.187		0.180		0.184	
Adj. R^2	−0.001		0.114		0.106		0.110	
Residual Std. Error	1.251		3.646		3.989		0.225	
F Statistic (df)	0.986 (9,100)		2.551* (9,100)		2.431* (9,100)		2.499* (9,100)	

Note. *** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$.

(Style B) vs. *explanation absence*. We included two different explanations in the design to run a robustness check for the results of Study 2. The two explanations were identical, except for the following: The *explanation Style B* presented the hotel characteristics according to the size of their marginal impacts (i.e., strong negative, weak negative, weak positive, strong positive), see Fig. C.13. (3) Following Task II, we measured participants' perceived task complexity. (4) At the end of the study, we included a manipulation check to verify that participants in the explanation presence condition were aware they had received an explanation. (5) We did not include the Study 2 measure for participants' perceived understanding of the explanation anymore.

Measures

This study included the same measures (for user perceptions and compliance) as Study 2, see also Table 2. In addition, the following measures were included:

Cognitive ability. We were looking for a measure that explicitly assessed the cognitive ability to handle complex visual tasks – due to the nature of the explanation provided in the experimental studies. Standardized psychological tests of cognitive ability typically measure core cognitive abilities – e.g., the Primary Mental Abilities (PMA) test (Guilford, 1972) or the Wonderlic Personnel Test (WPT) (McKelvie, 1989). Prior research indicates that individuals are capable of differentiating between distinct cognitive abilities when providing self-ratings (Jacobs & Roodenburg, 2014). We considered the visual processing dimension of the Cattell–Horn–Carroll (CHC) model (cf., Schneider and McGrew (2018)) the best fit for our requirements. The validated measure is designed to indicate cognitive functioning in the CHC ability areas of visual processing (Jacobs & Roodenburg, 2014). To ensure participants were attentive during the main task (of estimating hotel room prices), we designed the study to be 15 min in total. This time constraint did not allow us to include a more complex cognitive ability test. For example, tests such as the Primary Mental Abilities (PMA) test (Guilford, 1972) or the Wonderlic Personnel Test (WPT) (McKelvie, 1989) would have increased the study's length by up to 30 min (Román-González, Pérez-González, & Jiméanez-Fernández, 2017). Though there are shorter versions of cognitive ability tests, they still take 8–10 min and usually only the full version has been validated.

Participants were informed that "the statements below ask you to rate how easy or difficult you usually find it is to perform certain tasks. Please compare yourself to most people your age. There are no right or wrong answers; you just need to pick the response that best describes you". Participants had to evaluate four statements on a scale from '1–extremely difficult' to '7–extremely easy', see Table 5. All statements concerned the visual processing dimension of cognitive ability in handling complex tasks. We considered this dimension as most appropriate and relevant for our analysis, as the explanations were visualized in tables.

Perceived task complexity. Perceived task complexity was measured with four items on a scale from '1–I totally disagree' to '7–I totally agree'; see Table 5.

Control variables and manipulation check. This study included the same control variables as Study 2 (except participants' perceived understanding of the explanation, which we planned to assess only once).

Table 5
Measures of cognitive ability and perceived task complexity (Study 3).

Construct	Item	Statement	Source
Cognitive		Compared to most people my age, I usually find tasks requiring me to...	Jacobs and Roodenburg (2014)
Ability	CA1	...understand information presented in a visual format	
	CA2	...interpret visually displayed information	
	CA3	...imagine what an object would look like from a different angle	
	CA4	...mentally rotate three-dimensional images in my mind	
Perceived	PTC1	I found this to be a complex task.	Maynard and Hakel (1997)
Task	PTC2	This task was mentally demanding.	
Complexity	PTC3	This task required a lot of thought and problem-solving.	
	PTC4	I found this to be a challenging task.	

This study also included a manipulation check for the explanation presence condition, ensuring participants were aware of the provided explanation. In particular, we asked: “Please try to recall: When you estimated the room prices for the six hotels a second time, did you receive additional information (in red, blue, and gray), showing how the model arrived at the predicted price for each hotel?” Possible responses were: “yes” or “no”. Participants passed the manipulation check if they chose “yes” (“no”) and were indeed in the explanation presence (vs. explanation absence) condition, which included (did not include) the additional information.

Statistical specification

First, we tested for the effect of explanation presence on user perceptions (trust and understanding: H2a.1 and H2a.2) and compliance (intended and actual: H2b.1 and H2b.2). The analysis included the two explanation presence conditions (i.e., Style A and Style B) and served as a robustness check for Study 2 results. We ran a separate multiple regression analysis for each of the four predicted outcomes. Following the results of the robustness check, we combined the two explanation presence conditions (Style A and Style B). Hence, in the following analyses, we compared two experimental conditions (as in Study 2): *explanation presence* and *explanation absence*. Second, we tested for the indirect effect of explanation presence – through perceived task complexity – on user perceptions (trust and understanding: H3a.1 and H3a.2) and compliance (intended and actual compliance: H3b.1 and H3b.2), and for the moderating effect of cognitive ability for user perceptions (trust and understanding: H4a.1 and H4a.2) and compliance (intended and actual compliance: H4b.1 and H4b.2); see also Fig. 2. We ran a separate moderated mediation analysis (Hayes, 2014, model no. 5) for each of the four predicted outcomes.

In all analyses, we controlled for participants’ demographics. As in Study 2, we controlled for how well participants understood the explanation – or would have, if they had received it. Lastly, we controlled for the study ID (i.e., 1–6). We ran the study in 6 separate blocks to prevent order and anchoring effects; the 6 hotels had different prices, and the 13 hotel characteristics had different impacts on the price.

6.2. Results

Robustness check: Effect of explanation presence on user perceptions

Despite the yet further simplified design of the explanation (following the results of Study 2), providing an explanation still neither improved participants’ trust (explanation (Style A): $p = 0.770$; explanation (Style B): $p = 0.697$), nor their understanding (explanation (Style A): $p = 0.674$; explanation (Style B): $p = 0.177$); see Fig. 3 and Table 6. Hypotheses 2a.1 and 2a.2 were not supported. Hence, the effects found in Study 2 are robust to slight changes in the design of the provided explanation.

Robustness check: Effect of explanation presence on user compliance

Providing an explanation did not improve participants’ intended compliance for explanation (Style A) ($p = 0.241$). Providing an explanation (Style B) marginally reduced intended compliance ($p = 0.053$). Still, providing an explanation did not improve participants’ actual compliance (Style A: $p = 0.934$; Style B: $p = 0.172$); see Fig. 3 and Table 6. Hypotheses 2b.1 and 2b.2 were not supported. Overall, Study 2 findings were robust; the effect did not depend on the specific explanation provided.

Following the results of the robustness check, we combined the two explanation presence conditions. We present the comparison of an explanation presence and an explanation absence condition (just as in Study 2) for the testing of the Hypotheses 3a and 3b (mediation of perceived task complexity) and Hypotheses 4a and 4b (moderation of cognitive ability). Appendix B details the results regarding the initial comparison between the three conditions.

Task complexity as a mediator to the effect of explanation presence on user perceptions

Providing an explanation increased participants’ perceived task complexity ($p < 0.001$). This increased perceived task complexity (marginally) impaired participants’ trust ($p = 0.061$) and understanding ($p < 0.001$). Hypothesis 3a.1 was partly, and 3a.2 was fully supported. See Table 7 for the regression results.

Task complexity as a mediator to the effect of explanation presence on user compliance

Providing an explanation increased participants’ perceived task complexity ($p < 0.001$), and therefore impaired participants’ intended compliance ($p = 0.011$). Hypothesis 3b.1 was supported. Finally, we find that this increased perceived task complexity did not affect participants’ actual compliance ($p = 0.753$); Hypothesis 3b.2 was not supported. Table 7 presents the regression results.

Cognitive ability as a moderator to the effect of explanation presence on user perceptions

Despite the overall negative effect of providing an explanation on participants’ perceptions, cognitive ability (marginally) moderated the effect of explanation presence on trust ($p = 0.055$) and understanding ($p = 0.014$).

Though the effect of explanation presence on trust did not vary for moderate and high levels of cognitive ability ($p = 0.193$ and $p = 0.147$), it is in the proposed direction: When moving from moderate to high levels of cognitive ability, providing an explanation was associated

Table 6
Results: Robustness-check: Effect of explanation presence on user perceptions and compliance (Study 3).

	Outcomes							
	User perceptions				User compliance			
	Trust		Understanding		Intended		Actual	
	coeff	*se*	*coeff*	*se*	*coeff*	*se*	*coeff*	*se*
Constant	5.41***	0.42	5.77***	0.40	5.60***	0.41	0.58***	0.12
Explanation (Style A) presence	−0.05	0.17	−0.07	0.16	−0.19	0.16	0.00	0.05
Explanation (Style B) presence	−0.06	0.16	−0.21	0.16	−0.31†	0.16	−0.07	0.05
Gender	−0.20	0.14	−0.17	0.14	−0.35*	0.14	−0.01	0.04
Age	0.01	0.01	0.00	0.01	0.01	0.01	0.00	0.00
Education	−0.06	0.07	0.03	0.07	−0.02	0.07	0.02	0.02
Experience hospitality mgmt.	−0.02	0.04	0.04	0.04	−0.05	0.04	−0.03	0.01*
Experience revenue mgmt.	−0.09†	0.05	−0.07	0.05	−0.03	0.05	0.00	0.01
Experience revenue software	−0.47*	0.23	−0.25	0.22	−0.08	0.22	−0.07	0.06
Act. understand. explanation	0.06	0.19	−0.22	0.18	−0.43*	0.19	0.04	0.05
Study ID	*included*		*included*		*included*		*included*	
N	263		263		263		263	
R^2	0.043		0.048		0.093		0.050	
Adj. R^2	−0.011		−0.005		0.042		−0.003	
Residual Std. Error	0.930		0.985		2.051		0.088	
F Statistic (df)	0.803 (14,248)		0.901 (14,248)		1.822* (14,248)		0.939 (14,248)	

Note. *** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$.

Table 7
Results: Mediating effect of perceived task complexity and moderating effect of cognitive ability (Study 3).

	Mediator		Outcomes							
	Perceived		User perceptions				User compliance			
	Task complexity		Trust		Understanding		Intended		Actual	
	coeff	*se*	*coeff*	*se*	*coeff*	*se*	*coeff*	*se*	*coeff*	*se*
Constant	3.06***	0.61	5.85***	0.82	5.95***	0.72	5.53***	0.81	1.04***	0.24
Explanation presence	0.81***	0.17	−1.43†	0.77	−1.57*	0.68	−0.97	0.77	−0.49*	0.22
Perceived task complexity			−0.10†	0.15	−0.29***	0.05	−0.13*	0.05	0.00	0.02
Cognitive ability			−0.16	0.13	−0.09	0.01	−0.05	0.13	−0.10**	0.03
Cog. ability × expl. pres.			0.29†	0.15	0.33*	0.13	0.16	0.15	0.09*	0.04
Age	0.01	0.01	0.01	0.01	0.00	0.01	0.01	0.01	0.00	0.00
Gender	0.07	0.17	0.20	0.14	0.18	0.12	0.35*	0.14	−0.10	0.04
Education	−0.01	0.08	−0.07	0.07	0.01	0.06	−0.04	0.07	0.02	0.02
Experience hospitality mgmt.	0.06	0.05	−0.01	0.04	0.07†	0.04	−0.04	0.04	−0.02†	0.01
Experience revenue mgmt.	0.04	0.06	−0.09†	0.05	−0.07†	0.04	−0.05	0.05	0.00	0.01
Experience revenue software	−0.62*	0.27	0.41†	0.22	0.07	0.19	−0.05	0.21	0.06	0.06
Act. understand. explanation	−0.25	0.23	−0.05	0.18	0.15	0.16	0.40*	0.19	−0.02	0.05
Study ID	*included*		*included*		*included*		*included*		*included*	
N	263		263		263		263		263	
R^2	0.118		0.069		0.215		0.101		0.053	
Adj. R^2	0.087		0.020		0.174		0.054		0.004	
Residual Std. Error	1.665		1.118		0.870		1.107		0.093	
F Statistic (df)	3.761*** (9,253)		1.549 (12,250)		5.711*** (12,250)		2.342** (12,250)		1.157 (12,250)	

Note. *** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$; † $p < 0.10$.

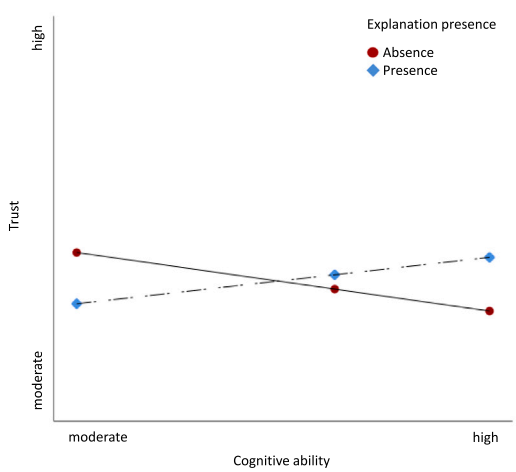


Fig. 4. Interaction of explanation presence and cognitive ability for user perceptions of trust (Study 3).

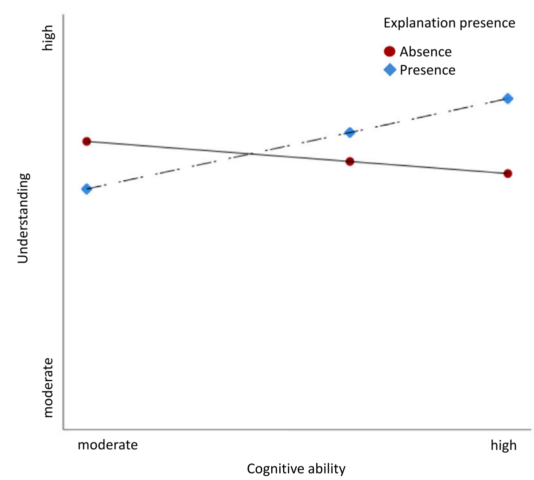


Fig. 5. Interaction of explanation presence and cognitive ability for user perceptions of understanding (Study 3).

with increasingly higher values of trust. In contrast, when moving from moderate to high levels of cognitive ability, not providing any explanation was associated with increasingly lower values of trust (see Fig. 4). Hence, Hypothesis 4a.1 was partly supported.

Next, if the cognitive ability was only moderate, providing an explanation did not improve participants' understanding ($p = 0.181$). Still, if the cognitive ability was high, providing an explanation improved participants' understanding ($p = 0.026$, see Fig. 5 and Table 7). Hypothesis 4a.2 was supported.

Cognitive ability as a moderator to the effect of explanation presence on user compliance

Participants' cognitive ability did not moderate the effect of explanation presence on intended compliance ($p = 0.270$). Hypothesis 4b.1 was not supported. However, participants' cognitive ability moderated the effect of explanation presence on actual compliance ($p = 0.035$). In particular, if the cognitive ability was only moderate, providing no explanation improved (rather than some explanation impaired) participants' actual compliance ($p = 0.044$) (see Fig. 6). If the cognitive ability was high, both providing and not providing an explanation yielded the same (moderately high) level of actual compliance ($p = 0.334$).

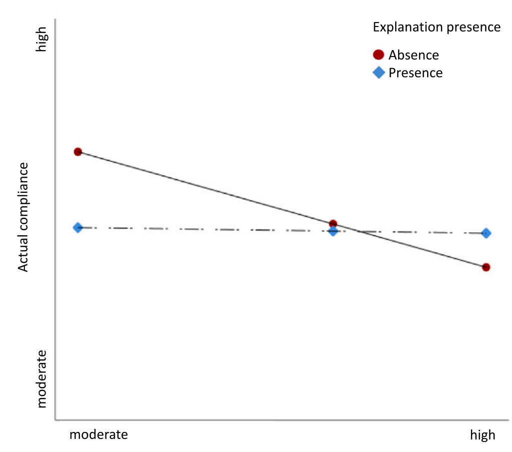


Fig. 6. Interaction of explanation presence and cognitive ability for user's actual compliance (Study 3).

Hypothesis 4b.2 was not supported. See Table 7 for the regression results.

6.3. Discussion

Study 3 confirmed that explanation presence neither improved user perceptions of trust and understanding nor user compliance. The results of Study 2 are robust to slight changes in the design of the provided explanation. Next, as proposed, we find that providing an explanation increased people's perceived task complexity, resulting in impaired user perceptions (of trust and understanding) and user (intended) compliance. Further, if an explanation was provided, perceptions of users with high cognitive ability were improved (for understanding – we see a very similar trend for trust). In contrast, (actual) compliance of users with high cognitive ability was not improved if an explanation was present. Still, actual compliance of users with low cognitive ability was impaired if an explanation was present.

7. General discussion

7.1. Summary of findings

This study aimed to find ways to address impaired user perceptions and low levels of compliance in human-AI collaboration. In three experimental studies, we focused on two techniques recognized in academic literature and widely used by practitioners: increasing users' decision control (the extent to which the AI's recommendation can be modified) and providing an explanation (detailing the impact each input feature had on the AI's recommendation). We learn that users benefit from enhanced decision control, while explanations present might harm user outcomes – unless they are appropriately designed for the specific user.

7.2. Theoretical contributions

We contribute to recent literature on human-AI collaboration by showing that enhanced decision control improves a variety of user outcomes. In particular, when users had more control over the task outcome, they trusted the AI system's recommendation more, and they felt they understood it better. In addition, high decision control increased user intended and actual compliance with the AI system's recommendation. Researchers have highlighted the need to give users decision control to satisfy their psychological needs and self-interest (Burton et al., 2020). Research on cognitive load theory (cf., Sweller, 1994) shows that enhancing control in a learning context yields high levels of task performance – while reducing extraneous and intrinsic load (inhibiting learning) and increasing germane load (fostering learning) (Van Merrienboer et al., 2002, 2002). We extend this theory to a modern learning context (also following (Hollender et al., 2010; Kirschner et al., 2011)) – where users encounter an AI system for the first time. Recent literature in similar contexts shows that decision control improves users' trust and task performance (e.g., when interacting with chatbots, see Benke et al. (2022)), and users' satisfaction with and the use of AI systems (e.g., when making forecasts, see Dietvorst et al. (2018)).

We also contribute to recent literature on human-AI collaboration by showing *why* and *when* explanations impair or improve user outcomes. Only approximately 5% of all papers studying AI explainability evaluate the quality of the provided explanations (Foerster, Klier, Kluge, & Sigler, 2020). While most work in the field relies on the researchers' intuition of what represents a "good" explanation (Miller, 2019), we test it in an experimental setting. In line with cognitive load theory, we show that providing explanations increases perceived task complexity (as a proxy for cognitive load), thus impairing user outcomes (Berthold et al., 2011). When users were provided with an explanation, they perceived the task as more complex and, in turn, both trusted and understood the AI system's recommendation less. Further, their intentions to comply with it were reduced. This aligns with previous work. For example, Zhao and Benbasat (2019) argued that transparency can hurt users' understanding and trust if excessively detailed information is provided. We identified a boundary condition to the negative effect of explanation presence on user outcomes. Cognitive load theory identified user expertise as an inhibitor of cognitive load, thus increasing task performance (Hollender et al., 2010; Van Gog & Ayres, 2009). We focused on cognitive ability, which plays a larger role than expertise when encountering a task for the first time (Kanfer & Ackerman, 1989). We find that the specific user's cognitive ability can compensate for the negative effect of providing explanations on user outcomes: Users' perceptions of understanding improved when an explanation was provided and the cognitive ability was high. We see a similar yet insignificant trend for perceptions of trust. In contrast, user compliance was reduced when an explanation was provided, and the cognitive ability was low. Previous researchers have proposed "user-centered explainable AI" (Ribera & Lapedriza, 2019) with the goal of providing the right quantity of high-quality information relevant to the specific user. We respond to this proposal by showing that the specific user's cognitive ability determines the effectiveness of a provided explanation.

In sum, we extend and integrate recent human-AI collaboration literature and cognitive load theory by investigating two techniques commonly used in human-AI collaboration (i.e., decision control and explanation presence) to improve various user outcomes.

7.3. Practical implications

Our work has direct implications for organizations. This setting is one of many possible settings where an AI system is trained and presented to users to improve task performance. The main idea is that the user should not make the decision alone because the task is too complex. In many scenarios, AI systems can, on average, handle complexity better than humans and provide a decision faster. Still, organizations often want to keep domain experts involved – either to monitor these systems or due to requirements by law. Today, decision control and explanations are frequently used as a means to address the multi-faceted challenges of AI adoption. We show that while organizations should generally provide users with enhanced decision control, they should handle explanations carefully. Providing explanations can have unintended consequences: Even if users objectively understand them, they might still be perceived as too complex, resulting in impaired rather than improved user outcomes. Our study identified three guidelines for organizations in designing effective explanations for users: First, "*Explanations may signal task complexity*". Organizations should take care that users understand the explanations and are not overwhelmed by the amount of information they provide. Otherwise, they will likely spend significantly more time on the task, and task outcomes will be impaired. Second, "*The cognitive ability of users needs to be taken into account*" – especially when the specific user has not yet encountered the task. Users will react more positively to a provided explanation and more likely act on it if it fits their cognitive ability. Third, "*The way information is communicated matters*". Overall, users react very similarly to explanations that convey the information in a slightly different way. Still, users might be confused if the information is sorted in a different way between tasks.

In practice, explanations are often designed by technical experts for non-technical users. Our study emphasizes that organizations need a multidisciplinary skill-set to ensure effective design and handling of explanations with regard to the individual user. One way could be to personalize explanations. Managers could consider the previous performance of the employees on complex tasks as an indicator of their cognitive ability. Alternatively, they could let employees choose independently whether they want to receive an explanation or not. We propose that employees who experience less cognitive overload when handling complex tasks would more likely want to receive additional information to help them solve the task.

7.4. Limitations and future research

Our work is not without limitations. First, we identify the following limitations regarding our theoretical foundation: Drawing from cognitive load theory (Sweller, 1994), and following Benke et al. (2022) and Van Merrienboer et al. (2002), we assume that enhanced decision control yielded positive user outcomes because it reduced extraneous and intrinsic load (i.e., hindering learning) and increased germane load (i.e., enhancing learning, see Vandewaetere and Clarebout (2013)). Future research should focus on the role of cognitive load in the effect of decision control on user outcomes and include the three different types of load in the study design. Similarly, following Berthold et al. (2011), we argue that providing explanations can impair user outcomes as cognitive load increases significantly. While we used perceived task complexity as a proxy for cognitive load, future research should replicate the effects using an explicit cognitive load measure (cf., Kirschner et al., 2011; Van Gog & Ayres, 2009).

Second, there are limitations inherent to the design of our study. Participants only had to make 6 consecutive decisions in each study. Compliance would be easier to judge with a larger number of tasks. If the number of tasks was increased, potential learning effects of participants' understanding of the provided explanation could be assessed as well. Further, we used a validated self-report measure to assess the participants' cognitive ability as outlined in Section 6.1 (Jacobs & Roodenburg, 2014). A wide variety of psychological constructs, such as intelligence, are assessed through self-report measures; these can be easily administrated and present a less anxiety-inducing assessment format Simms (2008). Still, there is a long and extensive history of the development and use of cognitive tests for psychological and HCI research. Hence, future research should investigate whether the study's findings can be replicated with classic, more extensive cognitive ability tests (Guilford, 1972; McKelvie, 1989). We suggest the use of validated self-report measures of cognitive ability in addition to – not instead of – classic cognitive ability tests.

Third, our work has limitations that open new avenues for future research and the management of human-AI collaboration. Research should continue to test how fine-grained details in the provided explanation affect user outcomes. We presented two very similar explanations to users. Still, we see slightly better effects for the explanation that sorted the model inputs arbitrarily but consistently for all subtasks rather than the one that sorted them by impact. We assume the latter explanation was slightly less effective because the order of model inputs was not consistent between the sub-tasks, following Shneiderman et al. Future research should also investigate a potential interaction effect of explanation presence and decision control in predicting collaboration outcomes (Shneiderman et al., 2016). Further, in this work, complying with the AI's recommendation means improved task performance – due to high system accuracy. We designed an AI system that incentivizes users to be compliant, and made sure that it indeed performed well before using it in the three studies (e.g., through hyperparameter tuning). In other settings, however, AI systems are less accurate, and compliance might indicate over-reliance – reducing task performance (Lee & See, 2004). Future research could vary both the accuracy of the AI recommendation and the task complexity, and test for people's appropriate reliance on the AI's recommendation. Bansal et al. (2021) show that providing explanations increased reliance on recommendations even when they were incorrect. Lastly, we suggest to seek replication of our findings in similar settings, where laypeople – instead of domain experts – make low-stakes decisions (e.g., Amazon purchases, Netflix movie selection). Our findings hint that laypeople would experience even higher levels of cognitive load when facing complex tasks.

7.5. Conclusions

This work investigates how enhancing decision control and providing explanations affects user perceptions of and compliance with AI-recommendations. Human-AI collaboration has become common, integrating highly complex AI systems – so called "black-box models" into the workplace. Recent literature shows that these highly complex AI systems potentially hamper user perceptions and compliance. In three experimental studies, we show that users and organizations generally benefit from enhancing users' decision control when collaborating with an AI system, while explanations may impair collaboration outcomes. We identify increased perceived task complexity as the underlying mechanism to this negative effect, and show that the specific user's cognitive ability determines the direction of this effect. This work bears both theoretical and practical implications for research in and management of human-AI collaboration; our study extends the integration of cognitive load theory into HCI research, also resolving inconsistent findings on the effectiveness of explanations in human-AI collaboration. And, our study gives clear guidance on how to manage human-AI collaboration, focusing in a variety of user perceptions and behaviors relevant to workplace efficiency and organizational performance.

CRediT authorship contribution statement

Monika Westphal: Conceptualization, Investigation, Data curation, Formal analysis, Writing – original draft, Writing – review & editing, Visualization, Project administration. **Michael Vössing:** Conceptualization, Methodology, Data curation, Formal analysis, Writing – original draft, Writing – review & editing, Funding acquisition. **Gerhard Satzger:** Conceptualization, Writing – review & editing, Funding acquisition, Resources, Supervision. **Galit B. Yom-Tov:** Conceptualization, Writing – review & editing, Supervision, Resources. **Anat Rafaeli:** Conceptualization, Writing – review & editing, Supervision, Resources.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgments

This research was supported by the Federal Ministry for Economic Affairs and Climate Action of Germany (BMWK) through the Smart Design and Construction project (Project ID: 01MK20016F), the Karlsruhe Institute of Technology, and the Technion–Israel Institute of Technology. We thank David Heigl for the support in the initial data curation and formal analysis. We also thank Wes Cowley for English language editing and proofreading.

Appendix A. Construct validation and descriptive statistics

A.1. Means, standard deviations and correlations

Tables A.8–A.10 detail the means and standard deviations for all study variables included in the theoretical model – in Study 1, 2 and 3, respectively. Pearson product-moment correlations determined the relationship between the various latent constructs. In Study 1, there was a strong, positive correlation between trust, understanding, and intended compliance. Actual compliance was strongly positively correlated with intended compliance. In Study 2, there was also a strong, positive correlation between trust, understanding, and intended compliance. In contrast to Study 1, actual compliance was not correlated with intended compliance. In Study 3, trust, understanding, and intended compliance were again strongly and positively correlated. In addition, cognitive ability was negatively correlated with perceived task complexity and positively correlated with understanding. Perceived task complexity was also negatively correlated with trust, understanding, and intended compliance. Actual compliance did not correlate with any of the constructs.

Table A.8
Means, standard deviations, and correlations (Study 1).

		Mean	SD	1	2	3	4
1	Trust	4.75	1.22	1			
2	Understanding	5.29	1.29	0.766**	1		
3	Intended Compliance	5.27	1.23	0.572**	0.564**	1	
4	Actual Compliance	0.38	0.33	0.140	0.140	0.287**	1

Notes. ** $p < 0.01$.

Table A.9
Means, standard deviations, and correlations (Study 2).

		Mean	SD	1	2	3	4
1	Trust	4.66	1.13	1			
2	Understanding	5.42	1.27	0.722**	1		
3	Intended Compliance	5.45	1.35	0.545**	0.640**	1	
4	Actual Compliance	0.46	0.32	0.115	−0.038	0.064	1

Notes. ** $p < 0.01$.

Table A.10
Means, standard deviations, and correlations (Study 3).

		Mean	SD	1	2	3	4	5	6
1	Cognitive Ability	5.08	0.98	1					
2	Perc. Task Complexity	3.62	1.35	−0.178**	1				
3	Trust	4.66	1.07	0.040	−0.138*	1			
4	Understanding	5.41	1.03	0.203**	−0.381**	0.603**	1		
5	Intended Compliance	5.55	1.08	0.067	−0.208**	0.598**	0.597**	1	
6	Actual Compliance	0.49	0.31	−0.091	−0.026	0.102	0.026	0.073	1

Notes. * $p < 0.05$, ** $p < 0.01$.

A.2. Scale reliabilities

Tables A.11–A.13 indicate that all Cronbach's alpha and composite reliability scores in Study 1 exceeded the threshold of 0.7, suggesting reliable internal consistency (Chin, 1998). Convergent validity was measured by the average variance extracted (AVE). All AVE values exceeded the recommended threshold of 0.5 (Fornell & Larcker, 1981) in all studies.

Table A.11
Reliability analysis (Study 1).

Construct	CA	CR	AVE
Trust	0.92	0.95	0.86
Understanding	0.89	0.93	0.82
Intended Compliance	0.91	0.96	0.92

Notes. CA = Cronbach's Alpha.
CR = Composite Reliability.
AVE = Average Variance Extracted.

Table A.12
Reliability analysis (Study 2).

Construct	CA	CR	AVE
Trust	0.99	0.94	0.84
Understanding	0.85	0.91	0.78
Intended Compliance	0.96	0.98	0.97

Notes. CA = Cronbach's Alpha.
CR = Composite Reliability.
AVE = Average Variance Extracted.

Table A.13
Reliability analysis (Study 3).

Construct	CA	CR	AVE
Cognitive Ability	0.79	0.87	0.62
Perc. Task Complexity	0.98	0.93	0.76
Trust	0.90	0.94	0.83
Understanding	0.79	0.88	0.70
Intended Compliance	0.90	0.95	0.91

Notes. CA = Cronbach's Alpha.
CR = Composite Reliability.
AVE = Average Variance Extracted.

A.3. Descriptives

Prior to hypothesis testing, we examined some descriptive statistics. First, the time participants spent on Task II (following the treatment): In Study 1, high decision control (marginally) decreased the time participants spent on the task by almost 1.5 min (M = 133 s, SD = 110 s, $p = 0.058$) compared to low decision control (M = 217 s, SD = 299 s). In Study 2, providing an explanation increased the time participants spent on the task by about half a minute (M = 169 s, SD = 108 s, $p = 0.041$) compared to if no explanation was present (M = 125 s, SD = 116 s). In Study 3, providing an explanation increased the time participants spent on the task, by more than one minute (explanation (Style A): M = 191 s, SD = 118 s, $p < 0.001$; explanation (Style B): M = 188 s, SD = 113 s, $p < 0.001$), compared to if the explanation was absent (M = 117 s, SD = 58 s).

In Study 2 and 3, we additionally examined participants' understanding of the explanation. In Study 2, 73% (80/110) of the participants objectively understood the explanation (provided in the end of the study), regardless of whether they were provided with an explanation or not in Task II. Compared to the pretest, where only 66% (50/76) understood the explanation (see 5.2) – participants' understanding increased by seven percentage points. In Study 3, participants' actual understanding of the explanation increased further. 85% (223/263) of the participants objectively understood the explanation, regardless of whether they were provided with it in the beginning of the study. This corresponds to an increase in participants' understanding by 12 percentage points compared to Study 2 and by 19 percentage points compared to the pretest for Study 2.

Appendix B. Supplementary analysis

Following the robustness check results for the direct effect of explanation presence on user outcomes (see Study 2), we combined the two explanation (Style A and Style B) presence conditions for further analysis (i.e., testing the indirect effect of explanation presence on user outcomes). In the following, we present the initial analysis that included the three experimental groups – explanation Style A, explanation Style B, no explanation – separately. We see that the two slightly different explanations behave quite similarly (see Figs. B.7–B.9 and Table B.14).

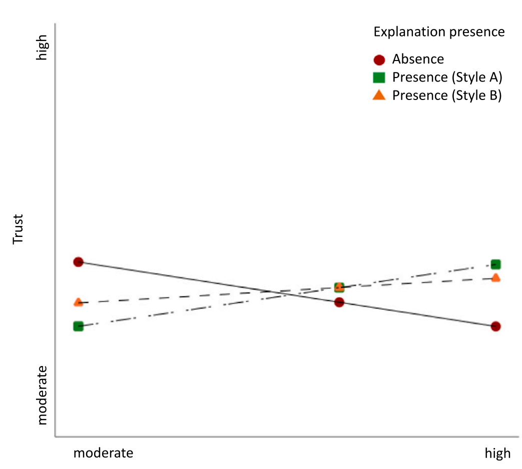


Fig. B.7. Interaction of explanation (Style A and Style B) presence and cognitive ability for user perceptions of trust (Study 3).

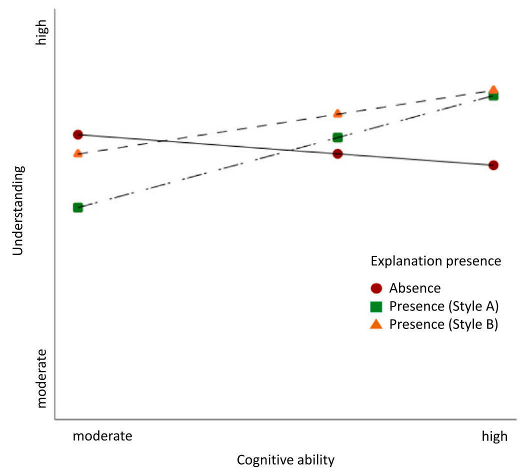


Fig. B.8. Interaction of explanation (Style A and Style B) presence and cognitive ability for user perceptions of understanding (Study 3).

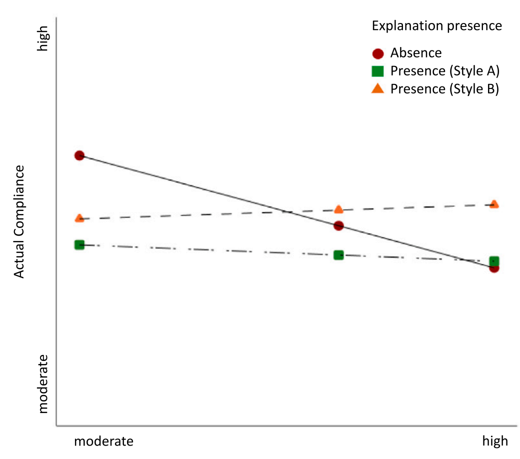


Fig. B.9. Interaction of explanation (Style A and Style B) presence and cognitive ability for user perceptions of compliance (Study 3).

Table B.14
Results Initial Analysis (Study 3): Mediating effect of perceived task complexity and moderating effect of cognitive ability.

	Mediator		Outcomes							
	Perceived		User perceptions				User compliance			
	Task complexity		Trust		Understanding		Intended		Actual	
	coeff	*se*	*coeff*	*se*	*coeff*	*se*	*coeff*	*se*	*coeff*	*se*
Constant	3.06***	0.61	5.83***	0.83	5.88***	0.73	5.53***	0.82	1.04***	0.24
Explanation (Style A) presence	0.82***	0.19	−1.63†	0.87	−1.96*	0.76	−0.93	0.86	−0.48†	0.25
Explanation (Style B) presence	0.81***	0.20	−1.20	0.90	−1.12	0.79	−1.03	0.89	−0.51*	0.26
Perceived task complexity			−0.10†	0.05	−0.28***	0.05	−0.14*	0.05	−0.01	0.02
Cognitive ability			−0.16	0.13	−0.09	0.11	−0.05	0.13	−0.10**	0.04
Cog. ability × expl. (Style A) pres.			0.33†	0.17	0.39**	0.15	0.15	0.17	0.08†	0.05
Cog. ability × expl. (Style B) pres.			0.24	0.17	0.25†	0.15	0.19	0.17	0.10*	0.05
Age	0.01	0.01	0.01	0.01	0.00	0.01	0.01	0.01	0.00	0.00
Gender	−0.07	0.17	0.20	0.14	0.19	0.12	0.36*	0.14	0.00	0.04
Education	−0.01	0.08	−0.06	0.07	0.02	0.06	−0.04	0.07	0.02	0.02
Experience hospitality mgmt.	0.60	0.05	−0.01	0.04	0.07†	0.04	−0.04	0.04	−0.03*	0.01
Experience revenue mgmt.	0.04	0.06	−0.09†	0.05	−0.07†	0.04	−0.04	0.04	0.00	0.01
Experience revenue software	−0.62*	0.27	0.41†	0.22	0.08	0.20	0.04	0.22	0.06	0.06
Act. understand. explanation	−0.25	0.23	−0.05	0.19	0.17	0.17	0.40*	0.19	−0.03	0.05
Study ID	*included*		*included*		*included*		*included*		*included*	
N	263		263		263		263		263	
R^2	0.118		0.070		0.221		0.103		0.062	
Adj. R^2	0.083		0.018		0.117		0.009		0.013	
Residual Std. Error	1.661		1.126		0.870		1.113		0.093	
F Statistic (df)	3.371*** (10,252)		1.337 (14,248)		5.038*** (14,248)		2.034* (14,248)		1.168 (14,248)	

Note. *** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$; † $p < 0.10$.

Appendix C. Study material

Figs. C.10–C.13 present screenshots of Task I (initial decision) and Task II (updated decision – either including no explanation, or explanation (Style A) or explanation (Style B)):

Your task now is to estimate the room price for six hotels in London.

The **average room price** across all hotels in London is **£180.**

Below are **13 hotel characteristics** that determine the room price for six specific hotels (Hotel ID: 1-6) in London**.**

Please estimate the room prices according to the information in the following table, and put the six estimations in the boxes below the table.

Hotel Characteristics (key facts, location, customer reviews)		Hotel 1	Hotel 2	Hotel 3	Hotel 4	Hotel 5	Hotel 6
Name	Scale	Value					
Breakfast Included	-	Yes	No	No	No	Yes	Yes
Number of Hotels within 1km	-	192	2	171	73	148	278
Number of Hotels in Brand	-	0	0	0	0	0	0
Number of Reviews	-	745	367	916	456	2142	2382
Distance to City Center in km	-	1.7	11.7	2.1	2.7	2.2	3.3
Number of Room Categories	-	5	14	3	7	4	5
Star Rating	0–5	3	0	3	5	4	2
Review Overall Score	1–10	8.8	8.1	7.2	8.7	7.7	7.9
Review Cleanliness Score	1–10	9.1	8.6	7.4	8.9	8.0	8.1
Review Comfort Score	1–10	8.4	7.9	6.5	8.6	7.4	7.3
Review Facilities Score	1–10	8.4	7.7	6.3	8.2	7.2	7.1
Review Location Score	1–10	9.4	7.7	8.6	9.4	8.7	9.0
Review Value for Money Score	1–10	8.3	7.9	6.6	7.8	6.9	7.8
Price of a Double Room for One Night [in £]							
Room Price		?	?	?	?	?	?

Fig. C.10. Screenshot of Task I: Initial decision (Studies 1–3).

Previously, you estimated the room price for six hotels based on the aforementioned 13 hotel characteristics. Remember that the **average room price** is **£180**.

We now use a machine learning model to predict the room price for specific hotels based on these 13 hotel characteristics.

You get the chance to **compare your estimations** from the previous task **with the model recommendations**: The tables below now also contain the model's **predicted** room price for each of the six hotels (Hotel ID: 1-6).

If you want, **you can change your previous estimations.**

Hotel 1 — Hotel Characteristics		
Name (key facts, location, customer reviews)	Scale	Value
Breakfast Included	-	Yes
Number of Hotels within 1km	-	192
Number of Hotels in Brand	-	0
Number of Reviews	-	745
Distance to City Center in km	-	1.7
Number of Room Categories	-	5
Star Rating	0–5	3
Review Overall Score	1–10	8.8
Review Cleanliness Score	1–10	9.1
Review Comfort Score	1–10	8.4
Review Facilities Score	1–10	8.4
Review Location Score	1–10	9.4
Review Value for Money Score	1–10	8.3
Price of a Double Room for One Night [in £]		
Predicted Room Price		**£217**
Room Price		?

Fig. C.11. Screenshot of Task II: Updated decision, explanation absence (Studies 1–3).

Previously, you estimated the room price for six hotels based on the aforementioned 13 hotel characteristics. Remember that the **average room price** is **£180**.

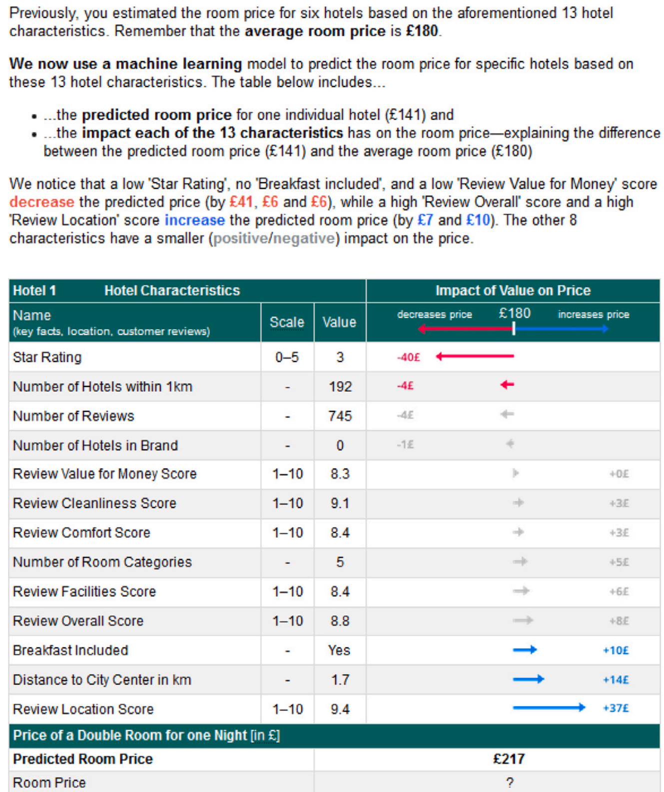
We now use a machine learning model to predict the room price for specific hotels based on these 13 hotel characteristics. The table below includes…

- …the **predicted room price** for one individual hotel (£141) and
- …the **impact each of the 13 characteristics** has on the room price—explaining the difference between the predicted room price (£141) and the average room price (£180)

We notice that a low 'Star Rating', no 'Breakfast included', and a low 'Review Value for Money' score **decrease** the predicted price (by £41, £6 and £6), while a high 'Review Overall' score and a high 'Review Location' score **increase** the predicted room price (by £7 and £10). The other 8 characteristics have a smaller (**positive/negative**) impact on the price.

Hotel 1 — Hotel Characteristics			Impact of Value on Price
Name (key facts, location, customer reviews)	Scale	Value	decreases price £180 increases price
Breakfast Included	-	Yes	+10£
Number of Hotels within 1km	-	192	-4£
Number of Hotels in Brand	-	0	-1£
Number of Reviews	-	745	-4£
Distance to City Center in km	-	1.7	+14£
Number of Room Categories	-	5	+5£
Star Rating	0–5	3	-40£
Review Overall Score	1–10	8.8	+8£
Review Cleanliness Score	1–10	9.1	+3£
Review Comfort Score	1–10	8.4	+3£
Review Facilities Score	1–10	8.4	+6£
Review Location Score	1–10	9.4	+37£
Review Value for Money Score	1–10	8.3	+0£
Price of a Double Room for one Night [in £]			
Predicted Room Price			**£217**
Room Price			?

Fig. C.12. Screenshot of Task II: Updated decision, explanation (Style A) presence (Study 2–3).


Previously, you estimated the room price for six hotels based on the aforementioned 13 hotel characteristics. Remember that the **average room price** is **£180**.

We now use a machine learning model to predict the room price for specific hotels based on these 13 hotel characteristics. The table below includes...

- ...the **predicted room price** for one individual hotel (£141) and
- ...the **impact each of the 13 characteristics** has on the room price—explaining the difference between the predicted room price (£141) and the average room price (£180)

We notice that a low 'Star Rating', no 'Breakfast included', and a low 'Review Value for Money' score decrease the predicted price (by £41, £6 and £6), while a high 'Review Overall' score and a high 'Review Location' score increase the predicted room price (by £7 and £10). The other 8 characteristics have a smaller (positive/negative) impact on the price.

Hotel 1 Hotel Characteristics			Impact of Value on Price
Name (key facts, location, customer reviews)	Scale	Value	decreases price £180 increases price
Star Rating	0–5	3	-40£
Number of Hotels within 1km	-	192	-4£
Number of Reviews	-	745	-4£
Number of Hotels in Brand	-	0	-1£
Review Value for Money Score	1–10	8.3	+0£
Review Cleanliness Score	1–10	9.1	+3£
Review Comfort Score	1–10	8.4	+3£
Number of Room Categories	-	5	+5£
Review Facilities Score	1–10	8.4	+6£
Review Overall Score	1–10	8.8	+8£
Breakfast Included	-	Yes	+10£
Distance to City Center in km	-	1.7	+14£
Review Location Score	1–10	9.4	+37£
Price of a Double Room for one Night [in £]			
Predicted Room Price			**£217**
Room Price			?

Fig. C.13. Screenshot of Task II: Updated decision, explanation (Style B) presence (Studies 3).

References

Adam, M., Wessel, M., & Benlian, A. (2021). AI-based chatbots in customer service and their effects on user compliance. *Electronic Markets*, *31*(2), 427–445. http://dx.doi.org/10.1007/s12525-020-00414-7.

Ahn, D., Almaatouq, A., Gulabani, M., & Hosanagar, K. (2021). Will we trust what we don't understand? Impact of model interpretability and outcome feedback on trust in AI. http://dx.doi.org/10.2139/ssrn.3964332.

Bansal, G., Wu, T., Zhou, J., Fok, R., Nushi, B., Kamar, E., Ribeiro, M. T., & Weld, D. S. (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* (pp. 1–16). http://dx.doi.org/10.1145/3411764.3445717.

Benke, I., Gnewuch, U., & Maedche, A. (2022). Understanding the impact of control levels over emotion-aware chatbots. *Computers in Human Behavior*, *129*, Article 107122. http://dx.doi.org/10.1016/j.chb.2021.107122.

Berthold, K., Röder, H., Knörzer, D., Kessler, W., & Renkl, A. (2011). The double-edged effects of explanation prompts. *Computers in Human Behavior*, *27*(1), 69–75. http://dx.doi.org/10.1016/j.chb.2010.05.025.

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

Campbell, D. T. (1988). Task complexity: A review and analysis. *Academy of Management Review*, *13*(1), 40–52. http://dx.doi.org/10.5465/amr.1988.4306775.

Cann, O. (2018). Machines will do more tasks than humans by 2025 but robot revolution will still create 58 million net new jobs in next five years. In *World economic forum*. https://www.weforum.org/press/2018/09/machines-will-do-more-tasks-than-humans-by-2025-but-robot-revolution-will-still-create-58-million-net-new-jobs-in-next-five-years/.

Chan, S. H., Song, Q., & Yao, L. J. (2015). The moderating roles of subjective (perceived) and objective task complexity in system use and performance. *Computers in Human Behavior*, *51*(Part A), 393–402. http://dx.doi.org/10.1016/j.chb.2015.04.059.

Chandrasekaran, A., Prabhu, V., Yadav, D., Chattopadhyay, P., & Parikh, D. (2018). Do explanations make VQA models more predictable to a human? In *Proceedings of the 2018 conference on empirical methods in natural language processing* (pp. 1036–1042). http://dx.doi.org/10.18653/v1/d18-1128.

Chin, W. W. (1998). The partial least squares approach to structural equation modeling. In G. A. Marcoulides (Ed.), *Modern methods for business research* (pp. 295–336). Lawrence Erlbaum Associates Publishers, http://dx.doi.org/10.4324/9781410604385.

Cialdini, R. B., & Goldstein, N. J. (2004). Social influence: Compliance and conformity. *Annual Review of Psychology*, *55*(1974), 591–621. http://dx.doi.org/10.1146/annurev.psych.55.090902.142015.

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. http://dx.doi.org/10.1037/xge0000033.

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

Ebel, P., Söllner, M., Leimeister, J. M., Crowston, K., & de Vreede, G. J. (2021). Hybrid intelligence in business networks. *Electronic Markets*, *31*(2), 313–318. http://dx.doi.org/10.1007/s12525-021-00481-4.

Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G* Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*(2), 175–191. http://dx.doi.org/10.3758/BF03193146.

Floridi, L., Cowls, J., Beltrametti, M., & ...,Vayena, E. (2018). AI4people—An ethical framework for a good AI society: Opportunities, risks, principles, and recommendations. *Minds and Machines*, *28*(4), 689–707. http://dx.doi.org/10.1007/s11023-018-9482-5.

Foerster, M., Klier, M., Kluge, K., & Sigler, I. (2020). Evaluating explainable artifical intelligence – what users really appreciate. In *Proceedings of the 28th European conference on information systems. Vol. 195*. https://aisel.aisnet.org/ecis2020_rp/195.

Fornell, C., & Larcker, D. F. (1981). Evaluating structural equation models with unobservable variables and measurement error. *Journal of Marketing Research*, *18*(1), 39. http://dx.doi.org/10.2307/3151312.

Giboney, J. S., Brown, S. A., Lowry, P. B., & Nunamaker, J. F. (2015). User acceptance of knowledge-based system recommendations: Explanations, arguments, and fit. *Decision Support Systems*, *72*, 1–10. http://dx.doi.org/10.1016/j.dss.2015.02.005.

Goddard, K., Roudsari, A., & Wyatt, J. C. (2012). Automation bias: A systematic review of frequency, effect mediators, and mitigators. *Journal of the American Medical Informatics Association*, *19*(1), 121–127. http://dx.doi.org/10.1136/amiajnl-2011-000089.

Grace, K., Salvatier, J., Dafoe, A., Zhang, B., & Evans, O. (2018). Viewpoint: When will ai exceed human performance? Evidence from ai experts. *Journal of Artificial Intelligence Research*, *62*, 729–754. http://dx.doi.org/10.1613/jair.1.11222.

Gregor, S., & Benbasat, I. (1999). Explanations from intelligent systems: Theoretical foundations and implications for practice. *MIS Quarterly: Management Information Systems*, *23*(4), 497–530. http://dx.doi.org/10.2307/249487.

Guilford, J. (1972). Thurstone's primary mental abilities and structure-of-intellect abilities. *Psychological Bulletin*, *77*(2), http://dx.doi.org/10.1037/h0032227.

Gutzwiller, R. S., & Reeder, J. (2021). Dancing with algorithms: Interaction creates greater preference and trust in machine-learned behavior. *Human Factors*, *63*(5), 854–867. http://dx.doi.org/10.1177/0018720820903893.

Haesevoets, T., De Cremer, D., Dierckx, K., & Van Hiel, A. (2021). Human-machine collaboration in managerial decision making. *Computers in Human Behavior*, *119*, Article 106730. http://dx.doi.org/10.1016/j.chb.2021.106730.

Hase, P., & Bansal, M. (2020). Evaluating explainable AI: Which algorithmic explanations help users predict model behavior? In *Proceedings of the 58th annual meeting of the association for computational linguistics* (pp. 5540–5552). http://dx.doi.org/10.18653/v1/2020.acl-main.491.

Hayes, A. F. (2014). Introduction to mediation, moderation, and conditional process analysis: A regression-based approach. *Journal of Educational Measurement*, *51*(3), 335–337. http://dx.doi.org/10.1111/jedm.12050.

Hemmer, P., Schemmer, M., Vössing, M., & Kühl, N. (2021). Human-AI complementarity in hybrid intelligence systems: A structured literature review. In *Proceedings of twenty-fifth pacific Asia conference on information systems (PACIS)* (pp. 1–14). https://www.researchgate.net/profile/Patrick-Hemmer-3/publication/352882174.

Hollender, N., Hofmann, C., Deneke, M., & Schmitz, B. (2010). Integrating cognitive load theory and concepts of human–computer interaction. *Computers in Human Behavior*, *26*(6), 1278–1288. http://dx.doi.org/10.1016/j.chb.2010.05.031.

Hsu, K.-C. (1993). *The effects of cognitive styles and interface designs on expert systems usage: An assessment of knowledge transfer* (Ph.D. thesis), Memphis State University, https://www.proquest.com/openview/7c5953efae23f19feab6ed4e7db7b28f/1?pq-origsite=gscholar&cbl=18750&diss=y.

Hyndman, R. J., & Koehler, A. B. (2006). Another look at measures of forecast accuracy. *International Journal of Forecasting*, *22*(4), 679–688. http://dx.doi.org/10.1016/j.ijforecast.2006.03.001.

Jacobs, K. E., & Roodenburg, J. (2014). The development and validation of the Self-Report Measure of Cognitive Abilities: A multitrait-multimethod study. *Intelligence*, *42*(1), 5–21. http://dx.doi.org/10.1016/j.intell.2013.09.004.

Jian, J., Bisantz, A., & Drury, C. (2000). Foundations for empirically determined scale of trust in automated systems. *International Journal of Cognitive Ergonomics*, *4*(1), 53–71, https://www.tandfonline.com/doi/abs/10.1207/S15327566IJCE0401_04.

Kanfer, R., & Ackerman, P. L. (1989). Motivation and cognitive abilities: An integrative/aptitude-treatment interaction approach to skill acquisition. *Journal of Applied Psychology*, *74*(4), 657. http://dx.doi.org/10.1037/0021-9010.74.4.657.

Kirschner, P. A., Ayres, P., & Chandler, P. (2011). Contemporary cognitive load theory research: The good, the bad and the ugly. *Computers in Human Behavior*, *27*(1), 99–105. http://dx.doi.org/10.1016/j.chb.2010.06.025.

Kizilcec, R. F. (2016). How much information? Effects of transparency on trust in an algorithmic interface. In *Proceedings of the 2016 CHI conference on human factors in computing systems*. http://dx.doi.org/10.1145/2858036.2858402.

Lee, M. K., Jain, A., Cha, H. J., Ojha, S., & Kusbit, D. (2019). Procedural justice in algorithmic fairness: Leveraging transparency and outcome control for fair algorithmic mediation. In *Proceedings of the ACM on human-computer interaction. Vol. 3* (CSCW), (pp. 1–26). ACM New York, NY, USA, http://dx.doi.org/10.1145/3359284.

Lee, J. D., & See, K. A. (2004). Trust in automation: Designing for appropriate reliance Human Factors. *Human Factors*, *46*(1), 50–80. http://dx.doi.org/10.1518/hfes.46.1.50_30392.

Legaspi, R., & Toyoizumi, T. (2019). A Bayesian psychophysics model of sense of agency. *Nature Communications*, *10*(4250), 1–11. http://dx.doi.org/10.1038/s41467-019-12170-0.

Liang, Y., Lee, S. A., & Jang, J. W. (2013). Mindlessness and gaining compliance in Computer-Human Interaction. *Computers in Human Behavior*, *29*(4), 1572–1579. http://dx.doi.org/10.1016/j.chb.2013.01.009.

Limerick, H., Coyle, D., & Moore, J. W. (2014). The experience of agency in human-computer interactions: A review. *Frontiers in Human Neuroscience*, *8*(643), 1–10. http://dx.doi.org/10.3389/fnhum.2014.00643.

Lind, E. A., Lissak, R. I., & Conlon, D. E. (1983). Decision control and process control effects on procedural fairness judgments. *Journal of Applied Social Psychology*, *13*(4), 338–350. http://dx.doi.org/10.1111/j.1559-1816.1983.tb01744.x.

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. http://dx.doi.org/10.1016/j.obhdp.2018.12.005.

Lundberg, S. M., Erion, G., Chen, H., DeGrave, A., Prutkin, J. M., Nair, B., Katz, R., Himmelfarb, J., Bansal, N., & Lee, S.-I. (2020). From local explanations to global understanding with explainable AI for trees. *Nature Machine Intelligence*, *2*(1), 56–67. http://dx.doi.org/10.1038/s42256-019-0138-9.

Lundberg, S. M., & Lee, S. I. (2017). A unified approach to interpreting model predictions. In *Proceedings of the 31st international conference on neural information processing systems* (pp. 4766–4775). https://proceedings.neurips.cc/paper/2017/file/8a20a8621978632d76c43dfd28b67767-Paper.pdf.

Mao, X., Sen, C., & Turner, C. (2018). A study in function modeling preferences and its variation with designer expertise and product types. In *International conference on-design computing and cognition* (pp. 631–648). Springer International Publishing, http://dx.doi.org/10.1007/978-3-030-05363-5.

Maynard, D. C., & Hakel, M. D. (1997). Effects of objective and subjective task complexity on performance. *Human Performance*, *10*(4), 303–330. http://dx.doi.org/10.1207/s15327043hup1004_1.

McKelvie, S. J. (1989). The Wonderlic Personnel Test: Reliability and validity in an academic setting. *Psychological Reports*, *65*(1), 161–162. http://dx.doi.org/10.2466/pr0.1989.65.1.161.

Miller, T. (2019). Explanation in artificial intelligence: Insights from the social sciences. *Artificial Intelligence*, *267*, 1–38. http://dx.doi.org/10.1016/j.artint.2018.07.007.

Mun, Y. Y., Jackson, J. D., Park, J. S., & Probst, J. C. (2006). Understanding information technology acceptance by individual professionals: Toward an integrative view. *Information & Management*, *43*(3), 350–363. http://dx.doi.org/10.1016/j.im.2005.08.006.

Norman, D. A., & Bobrow, D. G. (1975). On data-limited and resource-limited processes. *Cognitive Psychology*, *7*(1), 44–64, http://wexler.free.fr/library/files/norman%20(1975)%20on%20data-limited%20and%20resource-limited%20processes.pdf.

Nunes, I., & Jannach, D. (2017). A systematic review and taxonomy of explanations in decision support and recommender systems. *User Modeling and User-Adapted Interaction*, *27*(3–5), 393–444. http://dx.doi.org/10.1007/s11257-017-9195-0.

Nunes, I., Miles, S., Luck, M., & De Lucena, C. J. (2012). Investigating explanations to justify choice. In J. Masthoff, B. Mobasher, M. Desmarais, & R. Nkambou (Eds.), *User modeling, adaptation, and personalization* (pp. 212–224). Springer, http://dx.doi.org/10.1007/978-3-642-31454-4_18.

Obar, J. A., & Oeldorf-Hirsch, A. (2020). The biggest lie on the internet: Ignoring the privacy policies and terms of service policies of social networking services. *Information, Communication & Society*, *23*(1), 128–147. http://dx.doi.org/10.1080/1369118x.2018.1486870.

Paas, F., Tuovinen, J. E., Tabbers, H., & Van Gerven, P. W. (2003). Cognitive load measurement as a means to advance cognitive load theory. *Educational Psychologist*, *38*(1), 63–71. http://dx.doi.org/10.1207/S15326985EP3801_8.

Paas, F. G., & Van Merriënboer, J. J. (1994). Instructional control of cognitive load in the training of complex cognitive tasks. *Educational Psychology Review*, *6*(4), 351–371, https://link.springer.com/content/pdf/10.1007/BF02213420.pdf.

Papenmeier, A., Kern, D., Englebienne, G., & Seifert, C. (2022). It's complicated: The relationship between user trust, model accuracy and explanations in AI. *ACM Transactions on Computer-Human Interaction*, *29*(4), http://dx.doi.org/10.1145/3495013, https://doi.org/10.1145/3495013.

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. http://dx.doi.org/10.1109/3468.844354.

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

Peer, E., Brandimarte, L., Samat, S., & Acquisti, A. (2017). Beyond the Turk: Alternative platforms for crowdsourcing behavioral research. *Journal of Experimental Social Psychology*, *70*, 153–163. http://dx.doi.org/10.1016/j.jesp.2017.01.006.

Rader, E., Cotter, K., & Cho, J. (2018). Explanations as mechanisms for supporting algorithmic transparency. In *Conference on human factors in computing systems* (pp. 1–13). http://dx.doi.org/10.1145/3173574.3173677.

Renier, L. A., Schmid Mast, M., & Bekbergenova, A. (2021). To err is human, not algorithmic – Robust reactions to erring algorithms. *Computers in Human Behavior*, *124*, Article 106879. http://dx.doi.org/10.1016/j.chb.2021.106879.

Ribera, M., & Lapedriza, A. (2019). Can we do better explanations? A proposal of user-centered explainable AI. In *Joint proceedings of the ACM IUI 2019 workshop* (pp. 1–7). https://www.researchgate.net/profile/Mireia-Ribera-Turro/publication/339390078.

Román-González, M., Pérez-González, J.-C., & Jiméanez-Fernández, C. (2017). Which cognitive abilities underlie computational thinking? Criterion validity of the Computational Thinking Test. *Computers in Human Behavior*, *72*, 678–691. http://dx.doi.org/10.1016/j.chb.2016.08.047.

Schlicker, N., Langer, M., Ötting, S. K., Baum, K., König, C. J., & Wallach, D. (2021). What to expect from opening up 'black boxes'? Comparing perceptions of justice between human and automated agents. *Computers in Human Behavior*, *122*, Article 106837. http://dx.doi.org/10.1016/j.chb.2021.106837.

Schmidt, F. L. (2002). The role of general cognitive ability and job performance: Why there cannot be a debate. *Human Performance*, *15*(1–2), 187–210. http://dx.doi.org/10.1080/08959285.2002.9668091.

Schneider, W. J., & McGrew, K. S. (2018). The Cattell–Horn–Carroll theory of cognitive abilities. In D. Flanagan, & E. McDonough (Eds.), *Contemporary intellectual assessment: theories, tests, and issues* (pp. 73–163). The Guilford Press, https://psycnet.apa.org/record/2018-36604-003.

Schrader, C., & Bastiaens, T. J. (2012). The influence of virtual presence: Effects on experienced cognitive load and learning outcomes in educational computer games. *Computers in Human Behavior*, *28*(2), 648–658. http://dx.doi.org/10.1016/j.chb.2011.11.011.

Shin, D. (2020). How do users interact with algorithm recommender systems? The interaction of users, algorithms, and performance. *Computers in Human Behavior*, *109*, Article 106344. http://dx.doi.org/10.1016/j.chb.2020.106344.

Shin, D. (2021). The effects of explainability and causability on perception, trust, and acceptance: Implications for explainable AI. *International Journal of Human-Computer Studies*, *146*, Article 102551. http://dx.doi.org/10.1016/j.ijhcs.2020.102551.

Shin, D., & Park, Y. J. (2019). Role of fairness, accountability, and transparency in algorithmic affordance. *Computers in Human Behavior*, *98*, 277–284. http://dx.doi.org/10.1016/j.chb.2019.04.019.

Shneiderman, B., & Plaisant, C. (2006). Strategies for evaluating information visualization tools: Multi-dimensional in-depth long-term case studies. In *Proceedings of the 2006 AVI workshop on beyond time and errors: novel evaluation methods for information visualization* (pp. 1–7). http://dx.doi.org/10.1145/1168149.1168158.

Shneiderman, B., Plaisant, C., Cohen, M. S., Jacobs, S., Elmqvist, N., & Diakopoulos, N. (2016). *Designing the user interface: strategies for effective human-computer interaction. Vol. 6*. Pearson, https://www.orangeanalytics.com.au/sites/default/files/webform/designing-the-user-interface-strategies-for-effective-human-comp-ben-shneiderman-catherine-plaisant.pdf.

Simms, L. J. (2008). Classical and modern methods of psychological scale construction. *Social and Personality Psychology Compass*, *2*(1), 414–433. http://dx.doi.org/10.1111/j.1751-9004.2007.00044.x.

Skitka, L. J., Mosier, K., & Burdick, M. D. (2000). Accountability and automation bias. *International Journal of Human Computer Studies*, *52*(4), 701–717. http://dx.doi.org/10.1006/ijhc.1999.0349.

Skulmowski, A., & Xu, K. M. (2021). Understanding cognitive load in digital and online learning: A new perspective on extraneous cognitive load. *Educational Psychology Review*, 1–26. http://dx.doi.org/10.1007/s10648-021-09624-7.

Sweller, J. (1994). Cognitive load theory, learning difficulty, and instructional design. *Learning and Instruction*, *4*(4), 295–312. http://dx.doi.org/10.1016/0959-4752(94)90003-5.

Tariq, A., & Rafi, K. (2012). Intelligent decision support systems-A framework. *Information and Knowledge Management*, *2*(6), 12–20.

Thibaut, J., & Walker, L. (1978). A theory of procedure. *California Law Review*, *66*(3), 541. http://dx.doi.org/10.2307/3480099.

Tzafilkou, K., & Protogeros, N. (2017). Diagnosing user perception and acceptance using eye tracking in web-based end-user development. *Computers in Human Behavior*, *72*, 23–37. http://dx.doi.org/10.1016/j.chb.2017.02.035.

Van der Land, S., Schouten, A. P., Feldberg, F., van Den Hooff, B., & Huysman, M. (2013). Lost in space? Cognitive fit and cognitive load in 3D virtual environments. *Computers in Human Behavior*, *29*(3), 1054–1064. http://dx.doi.org/10.1016/j.chb.2012.09.006.

Van Gog, T., & Ayres, P. (2009). State of the art research into cognitive load theory. *Computers in Human Behavior*, *25*(2), 253–257. http://dx.doi.org/10.1016/j.chb.2008.12.007.

Van Merrienboer, J. J., Schuurman, J. G., De Croock, M., & Paas, F. (2002). Redirecting learners' attention during training: Effects on cognitive load, transfer test performance and training efficiency. *Learning and Instruction*, *12*(1), 11–37. http://dx.doi.org/10.1016/S0959-4752(01)00020-2.

Vandewaetere, M., & Clarebout, G. (2013). Cognitive load of learner control: Extraneous or germane load? *Education Research International*, *2013*, http://dx.doi.org/10.1155/2013/902809.

Venkatesh, V., Morris, M. G., Davis, G. B., & Davis, F. D. (2003). User acceptance of information technology: Toward a unified view. *MIS Quarterly*, *27*(3), 425–478. http://dx.doi.org/10.2307/30036540.

Vössing, M., Potthoff, F., Kühl, N., & Satzger, G. (2019). Designing useful transparency to improve process performance-evidence from an automated production line. In *Proceedings of the 27th European conference of information systems*. https://aisel.aisnet.org/ecis2019_rp/122.

Woznyj, H. M., Banks, G. C., Dunn, A. M., Berka, G., & Woehr, D. (2020). Re-introducing cognitive complexity: A meta-analysis and agenda for future research. *Human Performance*, *33*(1), 1–33. http://dx.doi.org/10.1080/08959285.2019.1689396.

Zhang, Y., Vera Liao, Q., & Bellamy, R. K. (2020). Effect of confidence and explanation on accuracy and trust calibration in AI-assisted decision making. In *FAT* 2020 - proceedings of the 2020 conference on fairness, accountability, and transparency* (pp. 295–305). http://dx.doi.org/10.1145/3351095.3372852.

Zhao, R., & Benbasat, I. (2019). Do users always want to know more? Investigating the relationship between system transparency and users' trust in advice-giving systems. In *Proceedings of the 27th European conference of information systems*. https://aisel.aisnet.org/ecis2019_rip/42.

Zhou, L., Paul, S., Demirkan, H., Yuan, L. I., Spohrer, J., Zhou, M., & Basu, J. (2021). Intelligence augmentation: Towards building human- machine symbiotic relationship. *AIS Transactions on Human-Computer Interaction*, *13*(2), 243–264. http://dx.doi.org/10.17705/1thci.00149.