

MASTERARBEIT | MASTER'S THESIS

Titel | Title

Navigating the Spectrum: Exploring Factors Influencing AI Aversion and
Appreciation Across Contexts

verfasst von | submitted by

Lara Derya Dranaz BSc

angestrebter akademischer Grad | in partial fulfilment of the requirements for the degree of
Master of Science (MSc)

Wien | Vienna, 2024

Studienkennzahl lt. Studienblatt |
Degree programme code as it appears on the
student record sheet:

UA 066 840

Studienrichtung lt. Studienblatt | Degree
programme as it appears on the student
record sheet:

Masterstudium Psychologie

Betreut von | Supervisor:

Univ.-Prof. Dipl.-Psych. Dr. Robert Böhm

Abstract

AI technology is advancing and becoming increasingly prominent in a diverse range of domains. As AI/algorithm aversion has become a widely shown phenomenon, it is crucial to gain a better understanding of the factors that contribute to it. In this preregistered study ($N = 98$), the influence of identity relevance and the ambiguity of the evaluative criteria on AI aversion was examined across contexts. Using mixed-effects regression models, the results indicate that both identity relevance and ambiguity significantly increase AI aversion, with an interaction effect suggesting a complex relationship between these factors. Further analysis suggests that positive experiences with AI mitigate aversion, while demographic factors such as age and education level also seem to affect AI perceptions. These findings emphasize the importance of context in the acceptance of AI and call for future research to further examine these effects in controlled settings and to identify additional influential factors to help develop strategies aimed at mitigating AI aversion.

Keywords: Algorithm aversion, identity relevance, ambiguity, human-AI interaction

Künstliche Intelligenz (KI) ist auf dem Vormarsch und gewinnt in verschiedensten Bereichen zunehmend an Bedeutung. Da KI-/Algorithmus-Aversion als weit verbreitetes Phänomen immer wieder gezeigt wurde, ist es wichtig, ein besseres Verständnis der Faktoren zu erlangen, die zu dieser Aversion beitragen. In dieser präregistrierten Studie ($N = 98$) wurde der Einfluss der Identitätsrelevanz und der Mehrdeutigkeit der Bewertungskriterien auf die KI-Aversion kontextübergreifend untersucht. Unter Verwendung von mixed-effects Regressionsmodellen zeigen die Ergebnisse, dass sowohl die Identitätsrelevanz als auch die Mehrdeutigkeit die KI-Aversion signifikant erhöhen, wobei ein Interaktionseffekt auf eine komplexe Beziehung zwischen diesen beiden Faktoren hindeutet. Weitere Analysen deuten darauf hin, dass positive Erfahrungen mit KI diese Aversion mindern, während auch demografische Faktoren wie Alter und Bildungsstand die Wahrnehmung von KI zu beeinflussen scheinen. Diese Ergebnisse unterstreichen die Bedeutung des Kontexts für die Akzeptanz von KI und verlangen nach zukünftiger Forschung, um diese Effekte in einem kontrollierten Umfeld weiter zu untersuchen, sowie zusätzliche Einflussfaktoren zu identifizieren, um bei der Entwicklung von Strategien zur Abschwächung der KI-Aversion zu helfen.

Navigating the Spectrum: Exploring Factors Influencing AI Aversion and Appreciation Across Contexts

Artificial intelligence (AI) is advancing to the point where tasks that were previously the exclusive domain of humans are now being carried out by machines. Therefore, AI is becoming increasingly prominent across a diverse range of domains, extending influence not only on major corporations but also on the everyday lives of individuals. The potential applications of AI are therefore numerous and diverse, with the capacity to enhance decision-making and task completion in a multitude of domains. Nevertheless, the advantages of AI can only be fully realized if it is accepted and employed by humans, and therefore further research in this area is indispensable.

Performance, perceptions, and adoption of AI (and of algorithms more generally) have been studied for many years. The phenomenon of *algorithm aversion* is particularly relevant here: It refers to the costly reluctance of individuals to follow algorithmic advice and their preference for human decision-makers over algorithmic ones, especially after seeing an algorithm err (Dietvorst et al., 2015). This phenomenon occurs even though algorithms frequently outperform humans in decision-making or forecasting tasks (Dawes et al., 1989; Elkins et al., 2013; Grove et al., 2000; Silver et al., 2017). Humans were even shown to be particularly averse to algorithms after observing them perform, despite witnessing the algorithms outperforming human forecasters (Dietvorst et al., 2015). One of the reasons for this is that individuals tend to lose trust in algorithms more quickly after witnessing an error (Dietvorst et al., 2015). Accordingly, humans appear to reject algorithms despite being familiar with their superior performance (Dietvorst et al., 2015).

Indeed, the previous literature shows many cases in which individuals reject superior AI/algorithms and exhibit aversion. Evidence indicates that human advice is not generally used more, but that in the case of errors or poor advice, the use of automated advice decreases more strongly compared to human advice (Prahl & Van Swol, 2017). The phenomenon of

AI/algorithm aversion has already been studied across cultures and its manifestations have been observed, for example, in the financial context (Liu et al., 2023). High aversion to algorithms was also shown in moral decision-making, which could prove challenging in a number of fields, including law and self-driving vehicles (Bigman & Gray, 2018). In addition, AI/algorithm aversion was observed in the medical field, as human physicians were accepted more likely (Wu et al., 2021) and individuals were less likely to use healthcare if the provider was automated rather than human (Longoni et al., 2019). It was also demonstrated that individuals tended to believe that they could better understand decisions made by human healthcare providers, once again making them more averse to algorithmic healthcare providers (Cadario et al., 2021). Further findings indicated that individuals were not opposed to the recommendations provided by AI, but rather to the AI advisor itself (Böhm et al., 2023). This is an important insight that highlights the necessity for a more in-depth examination of the perception of AI, rather than an analysis of the content itself.

Factors Influencing AI/Algorithm Aversion

Many theories and various factors that influence AI/algorithm aversion have already been analyzed, mostly concerning individual factors (such as psychological factors, motivation, demographic variables, familiarity with algorithms and tasks), algorithm factors (including understandability, ability to learn), and task factors (e.g. subjectivity) (Mahmud et al., 2022). Regarding the factors related to the algorithms themselves, it has been consistently shown that individuals tend to perceive algorithmic errors as being systematic, whereas they perceive human errors as being random and repairable (Dietvorst et al., 2015, 2018). It has furthermore been established that individuals are more inclined to rely on modifiable algorithms (Dietvorst et al., 2018). Along those lines, it was proposed that aversion could be reduced by emphasizing the ability of an algorithm to learn (Berger et al., 2021), as the avoidance of algorithmic advice frequently appears to be based on the assumption that these algorithms are incapable of learning

from their mistakes (Reich et al., 2023). The concept of ‘uniqueness neglect’ has been put forth as a similar driving factor for aversion to AI: In a medical context, the assumption that AI advisors do not adequately consider individual nuances and circumstances has been identified as a key factor contributing to algorithm aversion (Longoni et al., 2019). The concept of uniqueness neglect was demonstrated to be a significant factor in the aversion to AI once more in a financial context (Liu et al., 2023).

A factor that seems to affect an individual’s perception of AI is people’s prior experience with AI: individuals tend to question the competence of AI in generating high-quality outcomes but they are more likely to seek advice from an AI if they have had positive experiences with AI in the past (Böhm et al., 2023). Similarly, in one study consumers familiar with the use of an algorithm exhibited greater trust in algorithms than consumers who were unfamiliar with such technology (Castelo et al., 2019). Nevertheless, in another study, becoming acquainted with the advisor reduced reliance on the algorithmic advisor, but not on the human advisor, when participants saw that the advisors could make mistakes (Berger et al., 2021). Demonstrating the advisor’s ability to learn, however, served to neutralize this proposed adverse impact of familiarity (Berger et al., 2021), thereby providing an explanation for the contradictory results. In light of the effects of familiarity and experience with algorithms, it seems reasonable to suggest that the level of deliberate use of AI may also exert an influence, particularly as individuals gain conscious experience in such situations. It seems also possible that an individual’s technological affinity, and the potential knowledge and experience with algorithms that may accompany it, could influence the acceptance of AI, given the impact of familiarity, experience with algorithms, and the ability to learn. In relation to the individual itself, a potential impact of human self-confidence on the acceptance or rejection of AI recommendations has also been suggested: It is proposed that poor AI performance may have a negative effect on an individual’s self-confidence, leading not only to a loss of confidence in the AI itself when mistakes are observed, but also in the individual’s own abilities (Chong et

al., 2022). More fundamental characteristics of individuals, such as demographics, have been investigated as well, yet often fail to yield a coherent picture or direct connection (Park et al., 2021). With regard to gender, previous research has indicated that women tend to exhibit less trust in AI than men (Castelo & Ward, 2021). Additionally, older individuals have been observed to demonstrate lower levels of trust in AI (Castelo & Ward, 2021), although AI appreciation, on the other hand, has been found to be independent of age (Logg et al., 2019). Individuals who perceived themselves to be more unique additionally showed more aversion to medical AI (Longoni et al., 2019).

When looking at the tasks themselves, a higher level of algorithm aversion is evident in tasks that are perceived to be more subjective, as people tend to trust and rely on algorithms less in this case than in more objective tasks (Castelo et al., 2019). Higher levels of AI/algorithm aversion also hold true for tasks that are perceived as more consequential (Castelo & Ward, 2021). However, external circumstances also exert an influence: when individuals are pressed for time, they are less averse to algorithms, and willingly use them as an escape from the uncertainty and tension of the situation (Jung & Seiter, 2021).

AI/Algorithm Appreciation

Contrary to the findings of numerous studies on algorithm aversion, the opposite phenomenon is also evident: When individuals believed that the advice they were receiving came from an algorithmic source, they demonstrated a greater inclination to rely on it than when they thought that the advice came from another human (Logg et al., 2019). Furthermore, individuals seemed to exhibit a willingness to opt for algorithmic advice over their own judgment (Logg et al., 2019). This phenomenon has been defined as *algorithm appreciation* (Logg et al., 2019) and evidence of it has since been observed in other instances. For example, participants demonstrated a tendency to uncritically follow the recommendations of highly skilled AI agents when being advised in a game (Dunning et al., 2023). Moreover, a study that demonstrated the presence of algorithmic aversion in numerous instances found that, with

respect to specific tasks (including weather prediction, giving directions, and data analysis), participants exhibited a greater degree of trust in algorithms than in humans and not vice versa (Castelo et al., 2019).

Domain Dependence

The current literature on perceptions and use of algorithms thus demonstrates the existence of both aversion and appreciation towards AI/algorithms and has already begun to explore the potential influencing factors. It can be concluded that individuals exhibit both AI/algorithm aversion and (sometimes) AI/algorithm appreciation and that numerous factors appear to influence these attitudes. More recent studies have repeatedly suggested that the occurrence of AI aversion or AI appreciation may be dependent on the specific application context and its nature (Böhm et al., 2023; Castelo et al., 2019; Castelo & Ward, 2021; Morewedge, 2022). This also serves to explain the aforementioned study, which identified instances of both aversion to AI in some and appreciation of AI in other contexts (Castelo et al., 2019). Research so far has examined perceptions and use of algorithms across a wide range of domains, from medical recommendations/diagnoses/treatment (Cadario et al., 2021; Castelo et al., 2019; Longoni et al., 2019; Reich et al., 2023; Wu et al., 2021), financial advice (Castelo et al., 2019; Liu et al., 2023), to advice in games (Chong et al., 2022; Dunning et al., 2023) and moral decision making (Jauernig et al., 2022).

In order to potentially explain the different results and influencing factors observed in these studies, it is, therefore, necessary to consider the differences between the contexts and their nature. For instance, the earlier-mentioned consequentialness of a context could be a crucial factor. Furthermore, it is proposed that there may be a stronger preference for humans than for algorithms in domains that are relevant to their identity (Morewedge, 2022). This is based on the premise that people evaluate algorithms by comparing them with themselves and the social groups to which they belong and that domains that are central to their identity may evoke self-enhancing and self-protective biases and result in biased evaluation (Morewedge,

2022). This proposition that there is higher aversion to AI/algorithms in areas relevant to identity, and lower aversion, even appreciation, for AI/algorithm in areas with low identity relevance is already supported by preliminary evidence (Böhm et al., 2023; Logg et al., 2019). It has additionally been put forth that personal identity relevance may be more pronounced in areas that entail personal encounters (Böhm et al., 2023).

In a similar vein, it is also assumed that the ambiguity of the evaluative criteria of a decision or task also has an impact and leads to higher AI aversion (Morewedge, 2022). This assumption is based on the aforementioned impact of subjectivity—lower levels of AI/algorithm aversion are evident in tasks that are perceived to be more objective rather than subjective (Castelo et al., 2019)—and the observation that explaining the objective criteria behind an AI decision leads to lower AI aversion (Cadario et al., 2021). In light of these two constructs, Morewedge (2022) theorized that individuals would demonstrate a heightened preference for humans over algorithms in domains relevant to their identity, or when the evaluative criteria are ambiguous. In his theoretical framework, he even proposed that in domains irrelevant to their identity and with clear evaluation criteria, people should be indifferent between equivalent humans and algorithms and that people might even show algorithms appreciation in these domains (Morewedge, 2022). The nature of the analyzed domains regarding their identity relevance and the ambiguity of the evaluative criteria could therefore be a possible explanation for the different and partly inconsistent findings regarding AI aversion and appreciation in the literature to date.

The Present Research

This thesis draws upon Morewedge's (2022) theoretical framework and potential explanation, as well as the extensive research conducted on algorithm aversion in recent years. To take account of the rapid developments in the field of AI and its increasing integration into everyday life, the focus is set on AI aversion as opposed to algorithm aversion. Rather than considering only a single domain (e.g., the application of AI in medicine), factors influencing

AI aversion are examined across different contexts. A particular emphasis is placed on the two potentially explanatory constructs of identity relevance and ambiguity of the evaluative criteria as put forward theoretically by Morewedge (2022) but lacking any systematic investigation so far. Following this theoretical framework, we derived the following hypotheses:

H1: Participants will express higher levels of AI aversion in domains where the task is highly relevant to their identity compared to tasks with lower identity relevance.

H2: Participants will exhibit higher AI aversion in domains where the evaluative criteria of the task are more ambiguous compared to tasks with clearer criteria.

Additionally, the potential impact of technology affinity and the frequency of deliberate use of AI is explored, alongside the previously established influence of prior positive AI experiences.

Method

Design

The study used a within-participants design to investigate the factors influencing people's willingness to rely on AI to complete different tasks. We gave participants a short description of a specific task for each of the ten contexts in randomized order, one at a time. After each context, they were asked to answer the same set of items regarding their trust in an AI to complete the task, their trust in a human advisor, the ambiguity of the evaluative criteria, and their identity relevance of the context. The main goal was to examine how the ambiguity of the evaluative criteria and identity relevance of the context affect participants' levels of AI aversion/appreciation.

Sample

We conducted an a-priori power analysis for mixed-effects regressions using data simulation in RStudio (for further details/code, see OSF supplementary material), following the approach of DeBruine and Barr (2021). The power analysis indicated that a minimum

sample size of 85 will be required to achieve a power of 0.80 at an alpha level set at .05. Based on this analysis and to account for possible missing data, we preregistered to recruit at least 100 participants. The participants were recruited via the online recruitment platform Prolific (<https://www.prolific.com>). To participate in the study, individuals were required to speak English fluently, as the survey was conducted in English.

From the overall 105 participants recruited, we excluded seven participants who did not pass a simple attention check, as preregistered. Correspondingly, the final sample consists of $N = 98$ participants from a total of 22 countries (74 % Europe, 15% Africa, 3% South America, 2% Asia, 1% North America, 5% NA). Of those 98 participants 48 (49%) identified as female, 49 (50%) as male and one person chose not to disclose their gender identity (1%). Their mean age was $M = 29.6$ ($SD = 7.9$), ranging from 19 to 56 years and 69% of participants had at least a bachelor's degree. Most participants were white ($n = 73$), followed by participants of black or African American ($n = 13$), Asian ($n = 4$), and mixed ethnicity ($n = 3$). All participants were fluent in English and 26 % ($n = 25$) were native speakers.

Materials and Measures

The primary aim of presenting different contexts to the participants was to include contexts with varying levels of identity relevance and varying ambiguity in the evaluation criteria. When selecting the contexts, we therefore took care to cover a broad spectrum of areas and, to select contexts that varied in these terms. Following Morewedge's proposed matrix of ambiguity $\times$ identity relevance (Morewedge, 2022), contexts such as healthcare, writing, directions, financial advice, driving, and gift recommendations were selected. Additionally, some of the contexts were based on other previous studies: e.g., self-driving cars (Castelo et al., 2019; Jörling et al., 2019), financial advice (Liu et al., 2023), and medical diagnosis and treatment recommendations (Castelo et al., 2019; Castelo & Ward, 2021; Liu et al., 2023; Reich et al., 2023; Wu et al., 2021). We ensured that for all ten contexts and their described tasks, AI assistance was already conceivable and technically feasible. This led to the following ten

contexts: driving a car, writing a personal birthday letter to a spouse or close friend, career counseling, writing assistance/proofreading, giving directions, financial advice, gift recommendations for a spouse or close friend, therapeutic guidance via text, restaurant recommendations for a friend, and diagnosing diseases and providing treatment recommendations.

To give the participants a clear and consistent idea of the context, a specific task was described for each context. When characterizing the contexts, we ensured that each context described only one specific task and that the components and goals of the task were clearly stated. See Appendix A for all final contexts and descriptions provided to participants.

AI Aversion

After every description of the context, participants were asked how much they would trust an AI to perform the task on a scale from 1 (= *not trust at all*) to 100 (= *trust completely*) (based on Castelo & Ward, 2021). To assess AI aversion, they were then also asked how much they would trust a human advisor to perform the task on the same scale from 1 (= *not trust at all*) to 100 (= *trust completely*). AI aversion was then calculated as the difference between the level of trust in a human advisor minus the level of trust in the AI and could therefore range from -99 to 99.

Identity Relevance

To account for differences in the identity relevance of the contexts, participants then had to rate how relevant the task was to their personal values and identity after each context, ranging from 1 (= *not relevant at all*) to 7 (= *extremely relevant*).

Ambiguity

This was followed by an ad hoc item to assess ambiguity in the evaluative criteria of the task after every context, namely “To what extent do you perceive uncertainty in the criteria for achieving a correct outcome in this task?” measured on a 7-point Likert scale ranging from 1 (= *no uncertainty at all*) to 7 (= *high level of uncertainty*).

Technological Affinity

After participants had answered the items above for all ten contexts, the Affinity for Technology Interaction Short Scale (ATI-S) was used to measure technological affinity. It consisted of four items asking participants to rate statements describing their interactions with technical systems using a 6-point Likert scale ranging from 1 (= *completely disagree*) to 6 (= *completely agree*): “I like to occupy myself in greater detail with technical systems”, “I like testing the functions of new technical systems”, “It is enough for me that a technical system works; I don’t care how and why”, “It is enough for me to know the basic functions of a technical system” (Wessel et al., 2019). The ATI-S exhibited good internal consistency in this study, as indicated by a Cronbach’s alpha coefficient of $\alpha = .83$.

Frequency of Deliberate Use of AI

Subsequently, to account for differences in deliberate AI usage participants were asked to rate their approval to the statement “I frequently use AI in my daily life or work.” using a 7-point Likert scale ranging from 1 (= strongly disagree) to 7 (= strongly agree). The goal here was to measure the frequency of deliberate use of AI, as only incidents in which AI was used consciously are considered relevant here.

Positive Experience With AI

To account for differences in positive experience with AI, approval of the statement “I have had positive experiences with AI-generated content/advice.” was measured using a 7-point Likert scale ranging from 1 (= strongly disagree) to 7 (= strongly agree).

Demographic Information

Finally, we asked participants to indicate their age in years, their gender (female, male, non-binary, diverse, other), and their level of education (lower school leaving certificate/school leaving certificate without university entrance qualification - high school diploma or equivalent degree - Bachelor's degree - Master's degree/diploma - Doctoral degree - other).

Procedure

The study was a priori approved by the ethical review board of the Department of Occupational, Economic, and Social Psychology at the University of Vienna (DRB number: 2024/W/003) and informed consent was obtained from all participants. An online survey was developed using the SoSci survey software (<https://www.soscisurvey.de>) and all participants were recruited via Prolific. The survey could be completed on a personal computer or smartphone. Data collection was conducted on March 4th, 2024. The survey was estimated to take approximately 7 minutes to complete with an actual average processing time of 5:08 minutes ($SD = 1:22$ min). Financial compensation (hourly rate of approximately £7) for participation in the study was paid through the panel provider.

After giving consent, participants were presented with the first of ten contexts. Each context consisted of a description of a task and was followed by the assessment of AI aversion (trust in AI and trust in human advisor), identity relevance of the context, and ambiguity of the evaluative criteria. After completing all items in the first context, participants were given the next context and asked to complete the AI aversion, identity relevance, and ambiguity items again. Each participant was presented with all ten contexts and therefore answered subsequent items ten times. To account for habituation effects, we randomized the order in which participants completed the ten contexts. After answering all the items for all ten contexts, technological affinity, positive experience with AI, and frequency of deliberate use of AI were measured. At this point, we also included a simple item designed to check for attention. The item was presented in a straightforward manner, with the correct answer provided in the question itself. Finally, participants were asked to provide demographic information about their age, gender, and level of education.

Data Analysis

Variables and scales used were registered in advance. To perform the following analyses means were calculated for the frequency of deliberate use of and positive experience

with AI. Regarding AI aversion, identity relevance, and ambiguity, means were conducted for each of the ten contexts. The two reverse items of the ATI-S Scale (items 3 and 4) were recoded as required in the manual and then a mean was calculated for all four items (Wessel et al., 2019). Because each participant answered several items for each of the 10 contexts, data was converted from wide to long format. Consequently, 10 rows for each participant were created (one per context).

As preregistered, we then performed a mixed-effects regression since the observations were clustered within participants (given that each participant was assessed for all ten contexts). To investigate the two hypotheses presented, a random intercept model with subject as a random effect and identity relevance and ambiguity as fixed effects was conducted using the software program RStudio with the *lme4* package (Bates et al., 2015). The significance level (α) was set at .05. To meet preconditions for the statistical procedure, the dependent variable AI aversion was measured on a scale from -99 to 99 (as the two variables that we used to calculate the dependent variable were measured on a scale from 1 to 100), indicating a metric scale level. The normality of residuals was assessed using the Shapiro-Wilk test, which was found to be statistically significant, suggesting deviations from normality. However, a Q-Q plot of residuals indicated that the residuals were close to normally distributed with only minor deviations at the tails. The examination of a scatterplot of residuals revealed some clustering of points, indicating the potential for heteroscedasticity. We deemed it appropriate to proceed with the planned analysis, as mixed-effects models are generally robust to certain assumption violations, particularly with respect to the normality of residuals and homoscedasticity (Schielzeth et al., 2020). Variance Inflation Factor (VIF) values were close to 1 for both identity relevance and ambiguity, indicating no significant multicollinearity. The result of the Durbin-Watson test suggested no significant autocorrelation of residuals.

Furthermore, deviating from the preregistration, we carried out additional exploratory analyses using mixed-effects regressions: The potential interaction between identity relevance

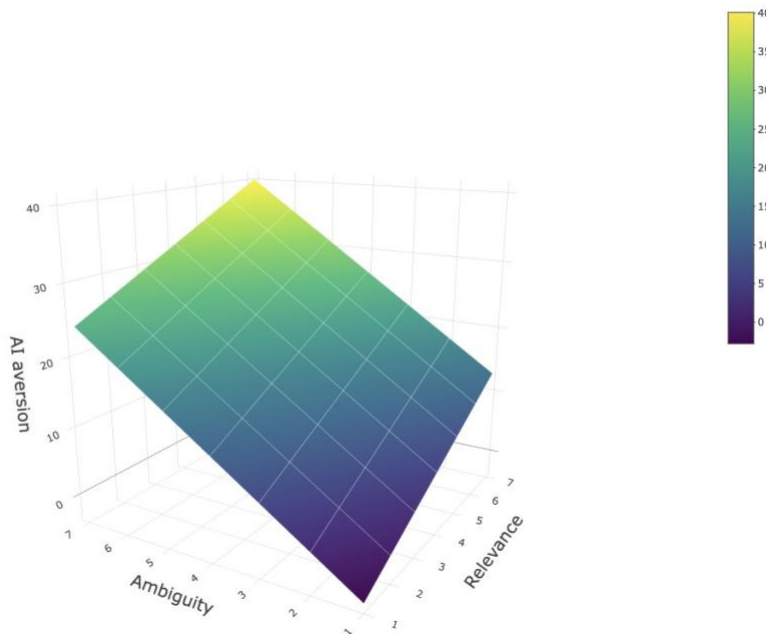
and ambiguity, as well as the roles of technological affinity, frequency of deliberate AI use, and positive experience with AI were investigated. Subsequently, we analyzed differences concerning gender, age, and level of education. Finally, the potential impact of fatigue was evaluated.

Results

All full models of analyses reported here, and additional robustness checks are reported in Appendix B. To first provide a foundational understanding of the primary predictors and their distributions, we calculated the mean and range for identity relevance and ambiguity as well as their correlation. The mean level of identity relevance was found to be 4.98, with a range from 1 to 7. The mean of ambiguity was 3.88, with a range from 1 to 7. The correlation between identity relevance and ambiguity was small ($r = .04$), suggesting a weak relationship within the dataset.

Confirmatory Analysis

The mixed-effects regression revealed significant positive main effects for both identity relevance ($t(837.89) = 5.74, p < .001, \beta = 0.20, 95\% \text{ CI } [0.13, 0.26]$) and ambiguity ($t(875.31) = 3.68, p < .001, \beta = 0.12, 95\% \text{ CI } [0.06, 0.19]$). This indicates that higher levels of identity relevance and ambiguity are associated with higher levels of AI aversion. The correlation between relevance and ambiguity was very low ($r = -0.05$). The analysis of random effects suggests that there is variability in the intercepts across different cases implying potential differences in the baseline AI Aversion among participants, with a variance of 154 ($SD = 12.41$) for the intercept. The intraclass correlation coefficient (ICC) was .139, suggesting that about 13.9% of the total variance in AI aversion is explained by these differences between the individuals. The mixed model explained 18.6% of the variance in AI aversion when considering both fixed and random effects (conditional $R^2 = .186$), whereas the fixed effects alone explained 5.4% of the variance (marginal $R^2 = .054$). Figure 1 illustrates the results of the confirmatory analysis.

Figure 1*3D-Surface Plot of AI Aversion*

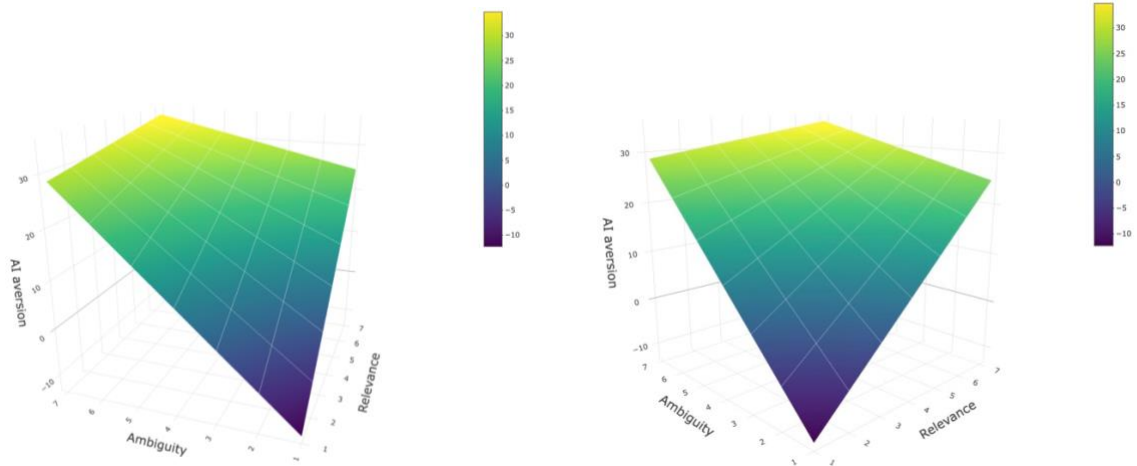
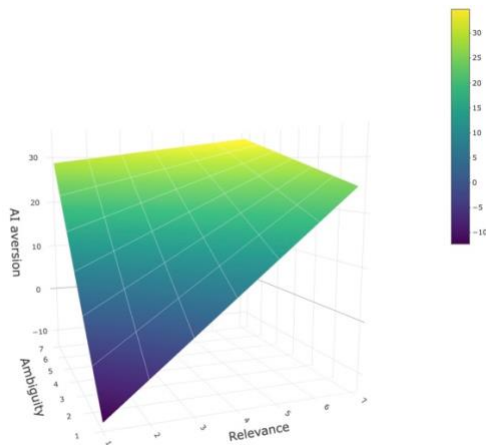
Note. 3D-Surface plot of AI aversion displaying the AI aversion score in relation to the ambiguity of the evaluative criteria as well as the identity relevance. Brighter colors and higher numbers indicate a higher level of AI aversion.

Exploratory Analysis

Interaction

In deviation from the preregistration, we subsequently included a potential interaction between identity relevance and ambiguity as a fixed effect in the mixed-effect model from the confirmatory analysis. The analysis again yielded significant positive main effects for both identity relevance ($t(938.06) = 4.53, p < .001, \beta = 0.19, 95\% \text{ CI } [0.12, 0.26]$) and ambiguity ($t(957.20) = 3.17, p = .002, \beta = 0.13, 95\% \text{ CI } [0.06, 0.20]$). This lends support to both hypotheses (H1 and H2). Additionally, the analysis of the interaction between identity relevance and ambiguity revealed a significant negative effect ($t(958.91) = -2.09, p = .037, \beta = -0.06, 95\% \text{ CI } [-0.12, 0.00]$) and suggests that the positive effect of identity relevance on AI aversion decreases as ambiguity increases and vice versa. In other words, when both identity relevance and ambiguity are high, their joint effect on AI aversion appears to be weaker than

would be expected if their effects were considered independently. The analysis of random effects once again indicated variability in the intercepts across different cases with a variance of 162.3 ($SD = 12.74$) for the intercept and suggested that about 14.7% of the total variance in AI aversion is explained by these differences between individuals. The mixed model with the interaction explained 19.8% of the variance in AI aversion when considering both fixed and random effects (conditional $R^2 = .198$), whereas the fixed effects alone explained 6.0% of the variance (marginal $R^2 = .060$). As the interaction explains additional variance, all subsequent exploratory analyses are based on this model with interaction term. Figure 2 illustrates the interaction effect, showing that as one of the two predictors increases, the influence of the other one on AI aversion diminishes.

Figure 2*3D Surface Plot of AI Aversion With Interaction Term**Angle a**Angle b**Angle c*

Note. 3D-Surface plot of AI aversion displaying the AI aversion score in relation to the ambiguity of the evaluative criteria as well as the identity relevance. Brighter colors and higher numbers indicate a higher level of AI aversion. To better illustrate the curvature of the graph, the same graph is presented here from three different angles (a–c).

Covariates

In a further exploratory analysis, we investigated the roles of technological affinity, frequency of deliberate AI use, and positive experience with AI. This was achieved by adding

these variables as additional fixed effects. Again, significant positive main effects were found for the factors identity relevance ($t(939.25) = 4.80, p = < .001, \beta = 0.20, 95\% \text{ CI } [0.13, 0.27]$) and ambiguity ($t(952.89) = 3.26, p = .001, \beta = 0.12, 95\% \text{ CI } [0.05, 0.19]$), as well as a significant interaction between the two ($t(955.87) = -2.26, p = .024, \beta = -0.06, 95\% \text{ CI } [-0.12, -0.01]$). Additionally, a significant negative effect was found for positive experiences with AI ($t(92.56) = -2.82, p = .006, \beta = -0.15, 95\% \text{ CI } [-0.25, -0.04]$), indicating that positive experiences are associated with lower AI aversion. No significant effects were found for technological affinity ($t(88.20) = -0.96, p = .342, \beta = -0.05, 95\% \text{ CI } [-0.14, 0.05]$) and the frequency of deliberate use of AI ($t(89.55) = 1.38, p = .170, \beta = 0.07, 95\% \text{ CI } [-0.03, 0.18]$). The analysis of random effects again demonstrated variability in the intercepts between the participants with a variance of 146.1 ($SD = 12.09$), suggesting that 13.4% of the total variance is explained by these differences between individuals, whereas fixed and random effects explained 19.9% and fixed effect alone explained 7.5% of the variance.

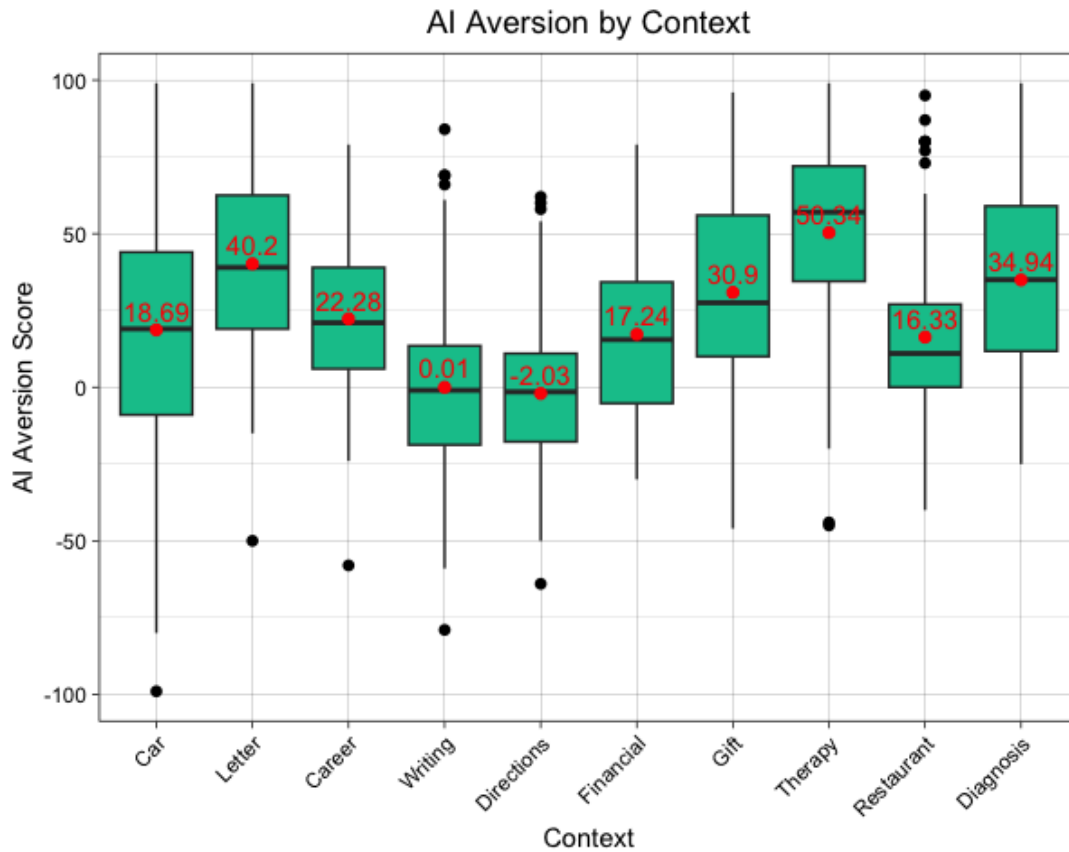
Demographic Analysis

We conducted a series of linear mixed-effects regression analyses with subsequent post-hoc tests (Tukey contrasts) to compare different groups in the demographic variables age, gender, and education. To account for the nested structure of the data, random intercept models with subject as random factor were once again used (detailed results can be found in Appendix B). When comparing AI aversion between female and male participants while controlling for relevance and ambiguity, no significant difference in AI aversion between male and female participants was found ($t(88.56) = -0.62, p = .538, \beta = -0.06, 95\% \text{ CI } [-0.25, 0.13]$). Participants aged 31 years and older showed significantly lower AI aversion compared to the youngest age group (18-25 years) when controlling for relevance and ambiguity ($t(86.65) = -3.61, p < .001, \beta = -0.39, 95\% \text{ CI } [-0.60, -0.18]$). The analysis of education levels and their association with AI aversion scores revealed significant differences between the lowest education level (lower school leaving certificate/school leaving certificate without university entrance qualification)

and all higher education levels (e.g., master's degree/diploma, doctoral degree). For instance, individuals with a master's degree or diploma exhibited on average lower AI aversion compared to participants with a lower school leaving certificate ($t(86.87) = -3.20, p = .002, \beta = -1.48, 95\% \text{ CI } [-2.39, -0.57]$). Similarly, significant differences were found between individuals holding a doctoral degree and those with a lower school leaving certificate ($t(86.28) = -3.07, p = .003, \beta = -1.69, 95\% \text{ CI } [-2.78, -0.61]$). No statistically significant differences in AI aversion were found between the different higher levels of education.

AI Aversion by Context

Figure 3 illustrates the distribution of AI aversion scores across the contexts presented, and indicates, that AI aversion varies by context: The contexts of therapy, diagnosis, and writing a letter to a spouse/close friend exhibit the highest mean AI aversion scores. In contrast, the contexts of writing assistance (in the sense of proofreading) and giving directions show the lowest mean scores. Some contexts, such as driving a car, display greater variability, indicated by a higher number of outliers. Overall, this suggests that AI aversion seems to be context-dependent in this dataset, with certain contexts seemingly eliciting stronger aversion to an AI than others.

Figure 3*Distribution of AI Aversion Across Contexts*

Note. Figure 3 illustrates the distribution of AI aversion across the various contexts, as displayed by box plots. The lower and upper fences represent the 25th and 75th percentiles, the line inside the box denotes the median value, and the mean values are indicated in red. Vertical extending lines show the values within the 1.5 interquartile range of the 25th and 75th percentile of each context. Filled circles indicate outliers.

Fatigue Effects

Finally, potential fatigue effects were evaluated due to the random order in which the contexts were presented. For this purpose, we calculated a variable containing the temporal position of each context presented and included it as a fixed effect in the mixed-effects model. This revealed a negative effect of context position ($t(858.27) = -2.50, p = .013, \beta = -0.07, 95\%$

CI [-0.13, -0.02]): As participants proceeded through more contexts (from the 1st to the 10th), their AI aversion decreased, indicating a potential fatigue effect.

Discussion

Given the increasing use of AI by laypeople and its growing integration into everyday life, it is important to better understand how people perceive and use AI. Therefore, the purpose of this study was to examine the influence of identity relevance and ambiguity on AI aversion across different contexts, following Morewedge's (2022) theoretical framework. In general, it can be stated that the results demonstrate a consistent aversion to AI across different cases, with the strength of this aversion varying considerably across contexts and individuals. The assumptions that participants would show higher levels of AI aversion in domains where the tasks were highly relevant to their identity or where the evaluative criteria of the task were more ambiguous could be confirmed. Both hypotheses were accepted, consequently. Hence, the results suggest that both identity relevance and ambiguity are important predictors of AI aversion, providing initial empirical support for Morewedge's (2022) assumptions. Specifically, contexts that participants find more relevant to their identity and situations with unclear evaluative criteria (higher ambiguity) were associated with a stronger aversion to AI. This provides a baseline understanding of the positive, individual effects of identity relevance and ambiguity on AI aversion. However, in deviation from the preregistration, we conducted an additional analysis to investigate the potential interaction between identity relevance and ambiguity. Despite the study not having been designed with this specific aim and the power analysis not accounting for this, a significant interaction effect was observed, even with a relatively small sample size. This significant interaction between identity relevance and ambiguity suggests a more complex dynamic where the combined impact of high relevance and high ambiguity on AI aversion appears to be less pronounced than their individual impacts would suggest: The effect of one predictor on AI aversion seems to depend on the level of the

other predictor, suggesting that as both relevance and ambiguity increase, their combined effect on AI aversion is less than the sum of their individual effects. The interaction highlights the importance of considering the combined effects of relevance and ambiguity on AI aversion rather than examining each predictor in isolation.

With respect to the additionally explored variables, it is crucial to highlight that previous positive experiences with AI were associated with lower levels of AI aversion. This finding is in line with prior research (Böhm et al., 2023) and contributes to the understanding of the mitigating effect of positive experience on AI aversion. Thus, the results support the implication proposed by Böhm et al. (2023) that providing individuals with opportunities to engage with and learn about AI could facilitate the acceptance and adoption of AI as an advisor. In addition, the effects of identity relevance and ambiguity could be utilized to better adapt AI to the needs and expectations of users and to foster greater acceptance. For instance, a reduction in the ambiguity of evaluative criteria (e.g. clearer guidelines for a “correct” result and transparent and understandable AI decision-making processes) and a reduction in the perceived relevance could help to mitigate AI aversion. In this way, the results can also guide policymakers and developers in prioritizing AI deployment in areas where it is likely to be more readily accepted, for example, in contexts where identity is less relevant. This could in turn result in more positive experiences with AI and, consequently, greater acceptance overall.

The variability in AI aversion scores across contexts suggests that acceptance of AI is in fact context-dependent and more pronounced in certain contexts than in others. The results of identity relevance and ambiguity suggest that these differences can be partially explained by the characteristics of the respective contexts. These findings emphasize the importance of taking contextual characteristics into account when assessing individuals’ attitudes toward AI. They illustrate the influence of contextual characteristics as a key explanatory factor in the different and occasionally contradictory findings regarding AI aversion and appreciation.

In contrast to previous findings, in this study, older participants (aged 31 and above) showed significantly less aversion to AI compared to the youngest participants (aged 18-25). Previous research has indicated that individuals of an older age cohort tend to exhibit a lesser degree of trust in AI technology than those in younger age groups (Castelo & Ward, 2021). It must be noted, however, that in the present study, the categories in question represent only a limited range of the typical age distribution of the general population well. The sample was notably young, with a mean age of approximately 30, which additionally necessitates a cautious interpretation of the results. The examination of the other age groups indicates that the age-related effects on AI aversion are not definitive and may be influenced by the relatively limited sample sizes in some age categories. The results therefore fail to provide further insight into the inconclusive earlier findings and highlight the necessity for a more comprehensive examination of the differences in AI aversion between age groups. Regarding gender, the results also do not support the previous hypothesis that women demonstrate less trust in AI than men (Castelo & Ward, 2021). Indeed, the present data revealed no significant difference between the genders in terms of AI aversion. On the contrary, a more coherent picture emerged around education. The findings indicate that individuals with higher levels of education tend to exhibit lower levels of AI aversion in comparison to those with lower levels of education. This implies a relationship between educational achievements and attitudes towards AI, indicating that higher education may shape perceptions of AI.

As the results indicate possible fatigue effects, it is important to consider why participants' AI aversion decreased over time (as they progressed through the contexts): The observed decline in AI aversion across successive contexts may be indicative of a reduction in participants' resistance to AI as a result of sustained exposure. This could mean that initial resistance to AI could be mitigated by prolonged exposure. However, an alternative explanation could be that the participants became less attentive over the course of the study. It is also possible that this decline in aversion may be attributed to the fact that with each

additional context seen, participants may have gained a more comprehensive understanding of the range of contexts represented in the study, and their anchoring of the values of trusting a lot and trusting less in AI may have shifted accordingly. Going forward, strategies should be explored to maintain participants' attention and accurately measure their attitudes toward AI throughout the duration of the study. This could include shorter sessions, varied task types, or regular breaks to reduce fatigue. Understanding and managing these effects is critical to the reliability of results.

Limitations and Future Research

Due to the extraction of the sample, the generalizability of these findings is limited as participants resided in 22 different countries, and interpretations should therefore be made with caution. However, it was ensured that all participants were fluent in English and the somewhat heterogeneous sample nevertheless provided an exciting first result.

Among the limitations of the study is furthermore the lack of manipulation and results must be viewed accordingly: This thesis employed Morewedge's (2022) theoretical framework in relation to AI aversion, considering the influence of identity relevance and ambiguity, and was able to demonstrate the value of examining these two variables. However, the variables were merely observed in different contexts, direct manipulation was not part of the study design and generalizability is therefore limited. In order to draw reliable conclusions, it is now necessary to extend the results of this study by directly manipulating the degree of ambiguity in the evaluative criteria and the identity relevance of the context. For instance, a subsequent study could examine a single context, varying its identity relevance and ambiguity in different conditions by manipulating additional information. Then participants could be randomly assigned to the conditions to retest the relationship between identity relevance and AI aversion, as well as ambiguity, in a more controlled manner. Moreover, it was beyond the scope of this study to further investigate the nature and origin of the interaction between relevance and ambiguity. Accordingly, further research is required to gain a more comprehensive

understanding of this interaction effect, particularly in contexts that are both highly relevant and highly ambiguous in their evaluative criteria.

As the analyses showed that the effect goes beyond ambiguity and identity relevance, future research could additionally revisit other context-specific variables that have already been discussed, such as perceived objectivity (Castelo et al., 2019) or consequentialness (Castelo & Ward, 2021) of a task. Additionally, it is important to continue examining individual-level factors, as the experience of the participants with AI has again been found to be a significant influencing factor.

Conclusion

In conclusion, this thesis emphasizes the significance of context in AI aversion, shedding light on the ways in which context-specific factors, namely identity relevance of the context and ambiguity of the evaluative criteria, appear to affect user perceptions. While the study design limits the generalizability of the results, this approach provides initial empirical support for the negative relationship between identity relevance and AI aversion, as well as ambiguity and AI aversion, as proposed by Morewedge (2022). These findings offer a more detailed insight into how individuals perceive and engage with AI, emphasizing the need for context-specific approaches in AI policy and development to enhance the acceptance and trust of users in AI as advisors. Future research should build upon this, investigating the relationship in greater detail within a controlled setting, and identifying additional critical characteristics of the contexts and the individuals. As AI continues to evolve, the ability to identify and understand these factors will become increasingly valuable in the design and implementation of AI aid.

Artificial Intelligence and Open Science Statement

The author would like to inform that in order to write this thesis the following Artificial Intelligence programs were used: ChatGPT, DeepL, and Grammarly. ChatGPT was mainly used for debugging R-Code, rephrasing, and finding a fitting title for the thesis. DeepL was used for translations between English and German, and Grammarly was used for proofreading.

Hypotheses and analyses were preregistered as of February 12, 2024. The preregistration as well as the study materials, data, and code used for analyzing the data are available on the Open Science Framework at <https://osf.io/6xyqu/>. All analyses were conducted with R (version 4.2.3).

References

- Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, 67(1). <https://doi.org/10.18637/jss.v067.i01>
- Berger, B., Adam, M., Rühr, A., & Benlian, A. (2021). Watch Me Improve—Algorithm Aversion and Demonstrating the Ability to Learn. *Business & Information Systems Engineering*, 63(1), 55–68. <https://doi.org/10.1007/s12599-020-00678-5>
- Bigman, Y. E., & Gray, K. (2018). People are averse to machines making moral decisions. *Cognition*, 181, 21–34. <https://doi.org/10.1016/j.cognition.2018.08.003>
- Böhm, R., Jörling, M., Reiter, L., & Fuchs, C. (2023). People devalue generative AI’s competence but not its advice in addressing societal and personal challenges. *Communications Psychology*, 1(1), 32. <https://doi.org/10.1038/s44271-023-00032-x>
- Cadario, R., Longoni, C., & Morewedge, C. K. (2021). Understanding, explaining, and utilizing medical artificial intelligence. *Nature Human Behaviour*, 5(12), 1636–1642. <https://doi.org/10.1038/s41562-021-01146-0>
- Castelo, N., Bos, M. W., & Lehmann, D. R. (2019). Task-Dependent Algorithm Aversion. *Journal of Marketing Research*, 56(5), 809–825. <https://doi.org/10.1177/0022243719851788>
- Castelo, N., & Ward, A. F. (2021). Conservatism predicts aversion to consequential Artificial Intelligence. *PLOS ONE*, 16(12), e0261467. <https://doi.org/10.1371/journal.pone.0261467>
- Chong, L., Zhang, G., Goucher-Lambert, K., Kotovsky, K., & Cagan, J. (2022). Human confidence in artificial intelligence and in themselves: The evolution and impact of confidence on adoption of AI advice. *Computers in Human Behavior*, 127, 107018. <https://doi.org/10.1016/j.chb.2021.107018>
- Dawes, R. M., Faust, D., & Meehl, P. E. (1989). Clinical Versus Actuarial Judgment. *Science, New Series*, 243(4899), 1668–1674. <https://doi.org/10.1126/science.2648573>
- DeBruine, L. M., & Barr, D. J. (2021). Understanding Mixed-Effects Models Through Data Simulation. 4(1). <https://doi.org/10.1177/2515245920965119>
- 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).

<https://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. <https://doi.org/10.1287/mnsc.2016.2643>

Dunning, R. E., Fischhoff, B., & Davis, A. L. (2023). When Do Humans Heed AI Agents' Advice? When Should They? *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 00187208231190459. <https://doi.org/10.1177/00187208231190459>

Elkins, A. C., Dunbar, N. E., Adame, B., & Nunamaker, J. F. (2013). Are Users Threatened by Credibility Assessment Systems? *Journal of Management Information Systems*, 29(4), 249–262. <https://doi.org/10.2753/MIS0742-1222290409>

Grove, W. M., Zald, D. H., Lebow, B. S., Snitz, B. E., & Nelson, C. (2000). Clinical versus mechanical prediction: A meta-analysis. *Psychological Assessment*, 12(1), 19–30. <https://doi.org/10.1037/1040-3590.12.1.19>

Jauernig, J., Uhl, M., & Walkowitz, G. (2022). People Prefer Moral Discretion to Algorithms: Algorithm Aversion Beyond Intransparency. *Philosophy & Technology*, 35(1), 2. <https://doi.org/10.1007/s13347-021-00495-y>

Jörling, M., Böhm, R., & Paluch, S. (2019). Service Robots: Drivers of Perceived Responsibility for Service Outcomes. *Journal of Service Research*, 22(4), 404–420. <https://doi.org/10.1177/1094670519842334>

Jung, M., & Seiter, M. (2021). Towards a better understanding on mitigating algorithm aversion in forecasting: An experimental study. *Journal of Management Control*, 32(4), 495–516. <https://doi.org/10.1007/s00187-021-00326-3>

Liu, N. T. Y., Kirshner, S. N., & Lim, E. T. K. (2023). Is algorithm aversion WEIRD? A cross-country comparison of individual-differences and algorithm aversion. *Journal of Retailing and Consumer Services*, 72, 103259. <https://doi.org/10.1016/j.jretconser.2023.103259>

Logg, J. M., Minson, J. A., & Moore, D. A. (2019). Algorithm appreciation: People prefer algorithmic to human judgment. *Organizational Behavior and Human Decision Processes*, 151, 90–103. <https://doi.org/10.1016/j.obhdp.2018.12.005>

- Longoni, C., Bonezzi, A., & Morewedge, C. K. (2019). Resistance to Medical Artificial Intelligence. *Journal of Consumer Research*, 46(4), 629–650. <https://doi.org/10.1093/jcr/ucz013>
- Mahmud, H., Islam, A. K. M. N., Ahmed, S. I., & Smolander, K. (2022). What influences algorithmic decision-making? A systematic literature review on algorithm aversion. *Technological Forecasting and Social Change*, 175, 121390. <https://doi.org/10.1016/j.techfore.2021.121390>
- Morewedge, C. K. (2022). Preference for human, not algorithm aversion. *Trends in Cognitive Sciences*, 26(10), 824–826. <https://doi.org/10.1016/j.tics.2022.07.007>
- Park, J., Hong, E., & Le, H. T. (2021). Adopting autonomous vehicles: The moderating effects of demographic variables. *Journal of Retailing and Consumer Services*, 63, 102687. <https://doi.org/10.1016/j.jretconser.2021.102687>
- Prahl, A., & Van Swol, L. (2017). Understanding algorithm aversion: When is advice from automation discounted? *Journal of Forecasting*, 36(6), 691–702. <https://doi.org/10.1002/for.2464>
- Reich, T., Kaju, A., & Maglio, S. J. (2023). How to overcome algorithm aversion: Learning from mistakes. *Journal of Consumer Psychology*, 33(2), 285–302. <https://doi.org/10.1002/jcpy.1313>
- Schielzeth, H., Dingemanse, N. J., Nakagawa, S., Westneat, D. F., Allegate, H., Teplitsky, C., Réale, D., Dochtermann, N. A., Garamszegi, L. Z., & Araya-Ajoy, Y. G. (2020). Robustness of linear mixed-effects models to violations of distributional assumptions. *Methods in Ecology and Evolution*, 11(9), 1141–1152. <https://doi.org/10.1111/2041-210X.13434>
- Silver, D., Schrittwieser, J., Simonyan, K., Antonoglou, I., Huang, A., Guez, A., Hubert, T., Baker, L., Lai, M., Bolton, A., Chen, Y., Lillicrap, T., Hui, F., Sifre, L., Van Den Driessche, G., Graepel, T., & Hassabis, D. (2017). Mastering the game of Go without human knowledge. *Nature*, 550(7676), 354–359. <https://doi.org/10.1038/nature24270>
- Wessel, D., Attig, C., & Franke, T. (2019). ATI-S - An Ultra-Short Scale for Assessing Affinity for Technology Interaction in User Studies. In *Proceedings of Mensch und Computer 2019* (pp. 147–154). ACM. <https://doi.org/10.1145/3340764.3340766>
- Wu, J., Xu, L., Yu, F., & Peng, K. (2021). Acceptance of Medical Treatment Regimens Provided by AI vs. Human. *Applied Sciences*, 12(1), 110. <https://doi.org/10.3390/app12010110>

Appendix A

Contexts

A) Driving Car

The task involves navigating a car through busy urban streets, handle traffic, and reach a destination seamlessly while ensuring a safe journey.

B) Write a Personal Birthday Letter to Spouse/Close Friend

The task entails composing a heartfelt letter to a spouse/close friend on their birthday, reminiscing about shared memories, and expressing love and appreciation.

C) Career Counseling

The task entails discussing professional aspirations, getting personalized career advice, and exploring potential career paths.

D) Writing Assistance

The task entails seeking writing assistance for a work report, to refine writing on grammar, punctuation, and spelling to ensure clarity and correctness.

E) Give Directions

The task entails providing clear and precise directions for a friend visiting the city for the first time, ensuring they reach a popular landmark.

F) Financial Advice

The task entails seeking financial advice for personal investment decisions, considering various factors such as risk tolerance and financial goals.

G) Gift Recommendations for Spouse/Close Friend

The task entails selecting a suitable present for a spouse/close friend, considering their diverse interests and preferences.

H) Therapeutic Chat

The task entails a scenario where someone seeks emotional support, expressing their feelings, and receiving therapeutic guidance via text.

I) Recommend Restaurant

The task entails suggesting a restaurant for a friend with open preferences for different cuisines and ambiances, catering to their varied tastes.

J) Diagnosing Diseases and Providing Treatment Recommendations

The task entails determining potential medical conditions and recommending suitable personalized treatment options based on the available information on symptoms.

Appendix B

Descriptives

Table 1

Sociodemographic Characteristics of Participants

Baseline characteristic						
	<i>n</i>	%	<i>M</i>	<i>SD</i>	min	max
Age	97		29.62	7.91	19	56
NA	1					
Gender						
Female	48	48.98				
Male	49	50.00				
NA	1	1.02				
Education						
Lower school leaving certificate / without university entrance qualification	1	1.02				
High school diploma or equivalent degree	29	29.59				
Bachelor's degree	43	43.88				
Master's degree / diploma	21	21.43				
Doctoral degree	2	2.04				
Other	2	2.04				
Country of Residence						
Portugal	20	20.41				
South Africa	15	15.31				
Poland	11	11.22				
United Kingdom	9	9.18				
Italy	8	8.16				
Greece	5	5.10				
Hungary	4	4.08				
Chile	3	3.06				
Belgium	2	2.04				
Estonia	2	2.04				
Spain	2	2.04				
Switzerland	2	2.04				
Canada	1	1.02				
Czech Republic	1	1.02				
Denmark	1	1.02				
France	1	1.02				
Germany	1	1.02				
Ireland	1	1.02				
Israel	1	1.02				
Lao	1	1.02				
Netherlands	1	1.02				
Slovakia	1	1.02				
NA	5	5.10				
Continent of Residence						
Europe	72	73.47				
Africa	15	15.31				
South America	3	3.06				

Asia	2	2.04
North America	1	1.02
NA	5	5.10
First Language		
English	25	25.51
Not English	71	72.45
NA	2	2.04
Ethnicity		
White	73	74.49
Black	13	13.27
Asian	4	4.08
Mixed	3	3.06
NA	5	5.10

Note. $N = 98$. Participants were on average 29.6 years old ($SD = 7.9$)

Confirmatory Analysis

Assumptions

Figure 4

Q-Q Plot of Residuals

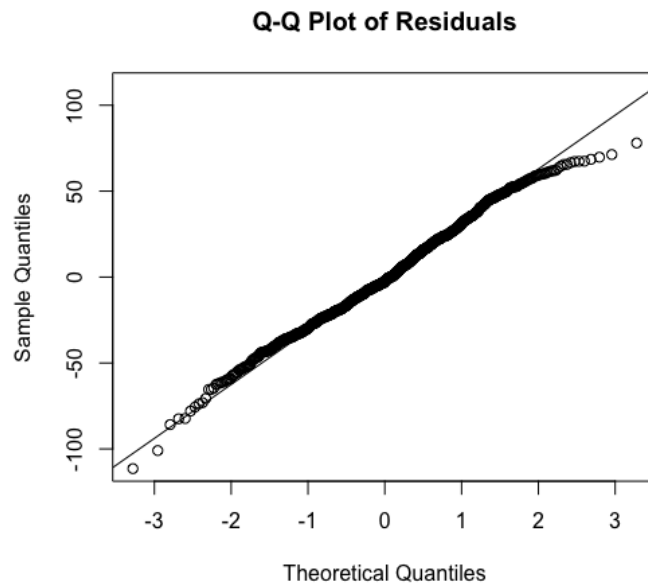


Figure 5

Scatterplot of Residuals

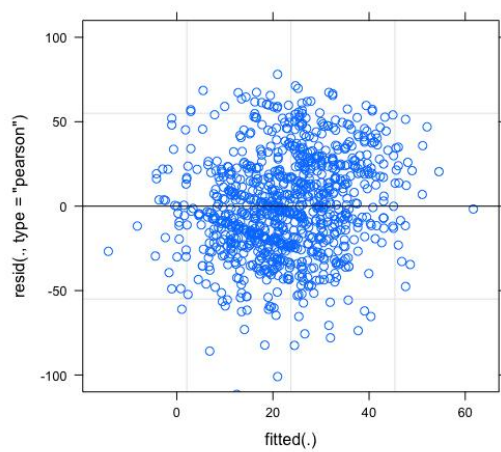


Table 2*Assumption Tests*

Variance Inflation Factor (VIF)	
Relevance	Ambiguity
1.002	1.002
Durbin-Watson test	
<i>d</i>	1.987
Shapiro-Wilk normality test	
<i>W</i>	0.995
<i>p</i> -value	.002**

* $p < .05$. ** $p < .01$. *** $p < .001$

Mixed-Effects Analysis**Table 3***Mixed-Effects Analysis: Relevance and Ambiguity*

Fixed effects										
	Estimate	<i>SE</i>	95% CI		<i>df</i>	<i>t</i> -value	<i>p</i> -value	Std. Beta	95% CI	
			(Estimate)						(Std. Beta)	
			LL	UL					LL	UL
Intercept	-9.94	4.95	-19.65	-0.23	685.50	-2.01	.045*	-.001	-.09	.09
Relevance	4.52	0.79	2.97	6.06	837.89	5.74	< .001***	.20	.13	.26
Ambiguity	2.64	0.72	1.23	4.04	875.31	3.68	< .001***	.12	.06	.19
Random effects										
			Variance		<i>SD</i>					
Participant (Intercept)			154		12.41					
Residual			952		30.85					
Model fit										
<i>R</i> ²		<i>R</i> ²		ICC						
(conditional)		(marginal)								
.186		.054		.139						
Correlation of Fixed effects										
			Intercept		Relevance					
Relevance			-.76		-					
Ambiguity			-.52		-.05					

Key: *p*-values for fixed effects calculated using Satterthwaite's approximations.

Confidence Intervals for estimates have been calculated using the Wald method.

Model equation: AI aversion ~ Relevance + Ambiguity + (1 | Participant)

Note. CI = confidence interval; LL = lower limit; UL = upper limit. **p* < .05. ***p* < .01. ****p* < .001

Exploratory Analysis

Interaction of Identity Relevance and Ambiguity

Table 4

Mixed-Effects Analysis: Relevance and Ambiguity and Interaction

Fixed effects										
	Estimate	SE	95% CI		df	t-value	p-value	Std. Beta	95% CI	
			(Estimate)						(Std. Beta)	
			LL	UL					LL	UL
Intercept	-26.17	9.11	-44.02	-8.31	901.39	-2.87	.004**	.001	-.09	.10
Relevance	7.71	1.70	4.37	11.05	938.06	4.53	< .001***	.19	.12	.26
Ambiguity	7.04	2.22	2.69	11.40	957.20	3.17	.002**	.13	.06	.20
Relevance x Ambiguity	-0.86	0.42	-1.68	-0.05	958.91	-2.09	.037*	-.06	-.12	.00
Random effects										
Variance			SD							
Participant (Intercept)	162.3		12.74							
Residual	945.2		30.74							
Model fit										
R ² (conditional)	R ² (marginal)		ICC							
.198	.060		.147							
Correlation of Fixed Effects										
	Intercept		Relevance		Ambiguity					
Relevance	-.94		-		-					
Ambiguity	-.89		.83		-					
Relevance x Ambiguity	.84		-.89		-.95					

Key: *p*-values for fixed effects calculated using Satterthwaite's approximations.

Confidence Intervals for estimates have been calculated using the Wald method.

Model equation: AI aversion ~ Relevance + Ambiguity + Relevance x Ambiguity + (1 | Participant)

Note. CI = confidence interval; LL = lower limit; UL = upper limit. **p* < .05. ***p* < .01. ****p* < .001

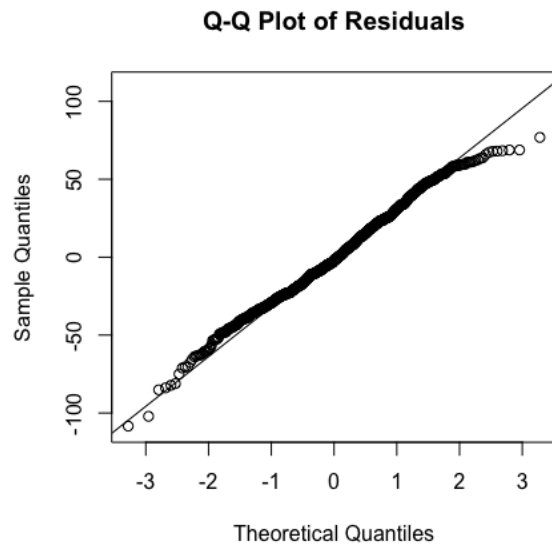
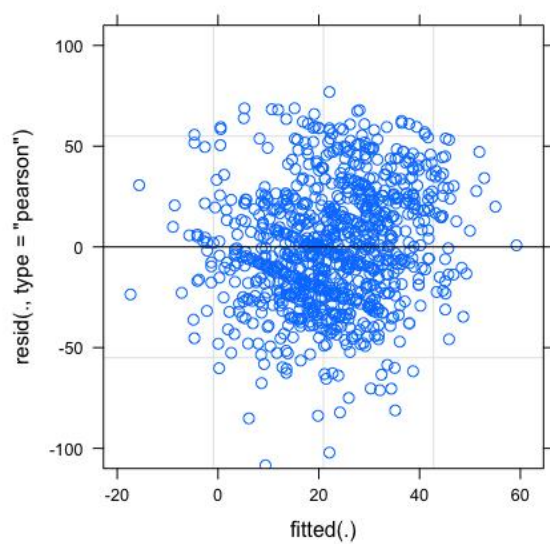
*Covariates***Figure 6***Q-Q Plot of Residuals***Figure 7***Scatterplot of Residuals*

Table 5*Assumption Tests*

Variance Inflation Factor (VIF)					
Relevance	Ambiguity	Technological Affinity	Frequency of use	Experience with AI	Relevance x Ambiguity
4.804	9.750	1.119	1.263	1.249	14.096
Durbin-Watson test					
<i>d</i>			1.986		
Shapiro-Wilk normality test					
<i>W</i>			0.994		
<i>p</i> -value			<.001***		

* $p < .05$. ** $p < .01$. *** $p < .001$

Table 6*Mixed-Effects Analysis: With Covariates*

Fixed effects										
	Estimate	SE	95% CI		df	t-value	p-value	Std.	95% CI	
			(Estimate)					Beta	(Std. Beta)	
			LL	UL					LL	UL
Intercept	-9.51	11.38	-31.81	12.79	324.17	-0.84	0.404	<.001	-.09	.09
Relevance	8.22	1.71	4.86	11.57	939.25	4.80	< .001***	.20	.13	.27
Ambiguity	7.23	2.22	2.88	11.59	952.89	3.26	.001**	.12	.05	.19
Relevance x Ambiguity	-0.94	0.41	-1.75	-0.13	955.87	-2.26	.024*	-.06	-.12	-.01
Technological Affinity	-1.47	1.54	-4.49	1.55	88.20	-0.96	0.342	-.05	-.14	.05
Frequency of deliberate AI use	1.34	0.97	-0.56	3.25	89.55	1.38	0.170	.07	-.03	.18
Experience with AI	-3.59	1.27	-6.08	-1.09	92.56	-2.82	0.006**	-.15	-.25	-.04
Random effects										
	Variance			SD						
Participant (Intercept)	146.1			12.09						
Residual	944.7			30.74						
Model fit										
R ²	R ²		ICC							
(conditional)	(marginal)									
.199	.075		.134							

Key: *p*-values for fixed effects calculated using Satterthwaite's approximations.

Confidence Intervals for estimates have been calculated using the Wald method.

Model equation: AIAversion ~ Relevance + Ambiguity + Relevance x Ambiguity + Technological Affinity + Frequency of deliberate AI use + Experience with AI + (1 | Participant)

Note. *N* = 98; CI = confidence interval; LL = lower limit; UL = upper limit. **p* < .05. ***p* < .01. ****p* < .001

Table 6b*Mixed-Effects Analysis: With Covariates - Correlation of Fixed Effects*

Variable	(Intr)	Relevance	Ambiguity	Technological Affinity	Frequency of deliberate AI use	Experience with AI
Relevance	-.71	-	-	-	-	-
Ambiguity	-.71	.83	-	-	-	-
Technological Affinity	-.38	-.02	.02	-	-	-
Frequency of deliberate AI use	-.08	.11	.06	-.20	-	-
Experience with AI	-.28	-.12	-.06	-.15	-.37	-
Relevance x Ambiguity	.66	-.89	-.95	< -.01	-.09	.08

Table 7*Mixed-Effects Analysis: With Gender*

Fixed effects										
	Estimate	SE	95% CI (Estimate)		df	t-value	p-value	Std. Beta	95% CI (Std. Beta)	
			LL	UL					LL	UL
Intercept	-24.92	9.26	-43.07	-6.77	849.28	-2.69	.007**	.03	-.10	.17
Relevance	7.73	1.71	4.37	11.08	927.27	4.52	< .001***	.20	.13	.26
Ambiguity	6.86	2.23	2.48	11.24	946.12	3.07	.002**	.13	.06	.19
Relevance x Ambiguity	-0.84	0.42	-1.66	-0.02	947.89	-2.02	.044*	-.06	-.11	.00
Gender (female)	- 2.04	3.30	-8.50	4.42	88.56	-0.62	.538	-.06	-.25	.13
Random effects										
	Variance		SD							
Participant (Intercept)	165.8		12.88							
Residual	949.7		30.82							
Model fit										
R^2 (conditional)	R^2 (marginal)		ICC							
.200	.061		.149							

Key: *p*-values for fixed effects calculated using Satterthwaite's approximations.

Confidence Intervals for estimates have been calculated using the Wald method.

Model equation: AI aversion ~ Relevance + Ambiguity + Relevance x Ambiguity + Gender + (1 | Participant)

Note. *n* = 97; CI = confidence interval; LL = lower limit; UL = upper limit. **p* < .05. ***p* < .01. ****p* < .001

Table 7b*Mixed-Effects Analysis: With Gender - Correlation of Fixed Effects*

Variable	(Intr)	Relevance	Ambiguity	Gender (Female)
Relevance	-.92	-	-	-
Ambiguity	-.87	.83	-	-
Gender (Female)	-.12	-.01	-.01	-
Relevance x Ambiguity	.83	-.89	-.95	< -.01

Table 8*Post-Hoc Tests for Gender*

Comparison	Estimate	SE	z-value	p-value
Female - Male	-2.04	3.30	-0.62	.536

* $p < .05$. ** $p < .01$. *** $p < .001$

Table 9*Mixed-Effects Analysis: With Age*

Fixed effects										
	Estimate	SE	95% CI		df	t-value	p-value	Std. Beta	95% CI	
			(Estimate)						(Std. Beta)	
			LL	UL					LL	UL
Intercept	-18.21	9.27	-36.38	-0.03	786.14	-1.96	.050	.19	.04	.35
Relevance	7.60	1.70	4.27	10.93	908.50	4.48	< .001***	.19	.12	.26
Ambiguity	6.75	2.22	2.40	11.10	939.44	3.04	.002**	.12	.05	.18
Relevance x Ambiguity	-0.85	0.41	-1.66	-0.04	944.84	-2.05	.041*	-.06	-.11	.00
Age group 26-30	-5.36	3.97	-13.14	2.41	86.16	-1.35	.180	-.16	-.39	.07
Age group 31+	-13.08	3.62	-20.19	-5.98	86.65	-3.61	< .001***	-.39	-.60	-.18
Random effects										
Variance			SD							
Participant (Intercept)	134.5		11.60							
Residual	950.7		30.83							
Model fit										
R ²	R ²		ICC							
(conditional)	(marginal)									
.197	.083		.124							

Key: *p*-values for fixed effects calculated using Satterthwaite's approximations.

Confidence Intervals for estimates have been calculated using the Wald method.

Model equation: AI aversion ~ Relevance + Ambiguity + Relevance x Ambiguity + Age + (1 | Participant)

Note. *n* = 97; reference group for age = 18-25 years; CI = confidence interval; LL = lower limit; UL = upper limit.**p* < .05. ***p* < .01. ****p* < .001

Table 9b*Mixed-Effects Analysis: With Age - Correlation of Fixed Effects*

Variable	(Intr)	Relevance	Ambiguity	Age group 26-30	Age group 31+
Relevance	-.91	-	-	-	-
Ambiguity	-.87	.83	-	-	-
Age group 26-30	-.19	.01	-.03	-	-
Age group 31+	-.19	-.03	-.02	.48	-
Relevance x Ambiguity	.82	-.89	-.95	-.03	.03

Table 10*Post-Hoc Tests for Age*

Comparison	Estimate	SE	z-value	p-value
Age 26-30 — age 18-25	-5.36	3.97	-1.35	.366
Age 31+ — age 18-25	-13.08	3.62	-3.61	< .001 ***
Age 31+ — age 26-30	-7.72	3.89	-1.98	.116

Note. Adjusted *p*-values reported: single-step method; **p* < .05. ***p* < .01. ****p* < .001

Table 11
Mixed-Effects Analysis: With Education

Fixed effects										
	Estimate	SE	95% CI (Estimate)		df	t-value	p-value	Std. Beta	95% CI (Std. Beta)	
			LL	UL					LL	UL
Intercept	17.24	16.67	-15.43	49.90	118.77	1.03	.303	1.36	0.47	2.25
Relevance	8.17	1.70	4.83	11.51	927.76	4.80	< .001***	0.19	0.12	0.26
Ambiguity	7.58	2.22	3.22	11.93	950.28	3.41	< .001***	0.13	0.06	0.20
Relevance x Ambiguity	-0.97	0.41	-1.79	-0.16	953.03	-2.35	.019*	-0.07	-0.12	-0.01
Education: High school diploma	-44.74	15.41	-74.95	-14.53	86.10	-2.90	.005 **	-1.33	-2.23	-0.43
Education: Bachelor’s degree	-43.57	15.36	-73.68	-13.46	86.72	-2.84	.006 **	-1.30	-2.19	-0.40
Education: Master’s degree / diploma	-49.75	15.55	-80.24	-19.27	86.87	-3.20	.002 **	-1.48	-2.39	-0.57
Education: Doctoral degree	-56.93	18.58	-93.34	-20.52	86.28	-3.07	.003 **	-1.69	-2.78	-0.61
Education: Other	-70.42	18.54	-106.77	-34.08	85.84	-3.80	< .001 ***	-2.10	-3.18	-1.01
Random effects										
	Variance		SD							
Participant (Intercept)	132.3		11.50							
Residual	944.8		30.74							
Model fit										
R ² (conditional)	R ² (marginal)		ICC							
0.199	0.087		0.123							

Key: p -values for fixed effects calculated using Satterthwaite's approximations.

Confidence Intervals for estimates have been calculated using the Wald method.

Model equation: AI aversion $\sim$ Relevance + Ambiguity + Relevance $\times$ Ambiguity + Education + (1 I Participant)

Note. $N = 98$; reference group for education = Lower school leaving certificate; CI = confidence interval;

LL = lower limit; UL = upper limit. * $p < .05$. ** $p < .01$. *** $p < .001$

Table 12*Mixed-Effects Analysis: With Education - Correlation of Fixed Effects*

Variable	(Intr)	Relevance	Ambiguity	High school diploma or equivalent	Bachelor's degree	Master's degree	Doctoral degree	Other
Relevance	-.41	-	-	-	-	-	-	-
Ambiguity	-.39	.84	-	-	-	-	-	-
High school diploma or equivalent	-.84	-.10	-.10	-	-	-	-	-
Bachelor's degree	-.83	-.11	-.11	.97	-	-	-	-
Master's degree	-.82	-.12	-.11	.96	.97	-	-	-
Doctoral degree	-.69	-.09	-.08	.80	.81	.80	-	-
Other	-.69	-.10	-.09	.80	.81	.80	.67	-
Relevance x Ambiguity	.38	-.89	-.95	.09	.09	.10	.06	.09

Table 13
Post-Hoc Tests for Education

Comparison	Estimate	SE	z-value	p-value
High school diploma – lower school leaving certificate	-44.74	15.41	-2.90	.033*
Bachelor's degree – lower school leaving certificate	-43.57	15.36	-2.84	.040*
Master's degree/diploma – lower school leaving certificate	-49.75	15.55	-3.20	.013*
Doctoral degree – lower school leaving certificate	-56.93	18.58	-3.07	.020*
Other – lower school leaving certificate	-70.42	18.54	-3.80	.002**
Bachelor's degree – High school diploma	1.18	3.65	0.32	.999
Master's degree/diploma – High school diploma	-5.01	4.34	-1.15	.830
Doctoral degree – High school diploma	-12.18	11.05	-1.10	.856
Other – High school diploma	-25.68	11.04	-2.33	.151
Master's degree/diploma – Bachelor's degree	-6.18	4.02	-1.54	.591
Doctoral degree – Bachelor's degree	-13.36	10.91	-1.22	.793
Other – Bachelor's degree	-26.85	10.93	-2.46	.111
Doctoral degree – Master's degree/diploma	-7.18	11.17	-0.64	.984
Other – Master's degree /diploma	-20.67	11.17	-1.85	.382
Other – Doctoral degree	-13.49	15.12	-0.90	.936

* $p < .05$. ** $p < .01$. *** $p < .001$

Table 14*Fatigue Effect: Mixed-Effects Analysis With Context Position*

Fixed effects										
	Estimate	SE	95% CI		df	t-value	p-value	Std. Beta	95% CI	
			(Estimate)						(Std. Beta)	
			LL	UL					LL	UL
Intercept	-20.66	9.35	-38.98	-2.34	910.57	-2.21	.027*	< .001	-.09	.10
Relevance	7.62	1.70	4.29	10.95	937.14	4.48	< .001***	.19	.12	.26
Ambiguity	6.86	2.22	2.51	11.21	956.17	3.09	< .002**	.13	.06	.19
Relevance x Ambiguity	-0.84	0.41	-1.65	-0.03	957.91	-2.04	.042*	-.06	-.11	.00
Context position	-0.86	0.34	-1.54	-0.19	858.27	-2.50	.013*	-.07	-.13	-.02
Random effects										
	Variance		SD							
Participant (Intercept)	161.3		12.70							
Residual	940.1		30.66							
Model fit										
R^2	R^2		ICC							
(conditional)	(marginal)									
.202	.065		.146							
Key: <i>p</i> -values for fixed effects calculated using Satterthwaite's approximations.										
Confidence Intervals for estimates have been calculated using the Wald method.										
Model equation: AI aversion ~ Relevance + Ambiguity + Relevance x Ambiguity + Context position + (1 Participant)										

Note. *N* = 98; CI = confidence interval; LL = lower limit; UL = upper limit. **p* < .05. ***p* < .01. ****p* < .001

Table 15*Fatigue Effect: Mixed-Effects Analysis With Context Position – Correlation of Fixed Effects*

Variable	(Intr)	Relevance	Ambiguity	Context position
Relevance	-.91	-	-	-
Ambiguity	-.87	-.83	-	-
Context position	-.24	.02	.03	-
Relevance x Ambiguity	.82	-.89	-.95	-.02