*Original Research Article*

BIG DATA & SOCIETY

Big Data & Society
October–December: 1–18

DOI: 10.1177/20539517241290217
journals.sagepub.com/home/bds
Sage

# Trade-offs in AI assistant choice: Do consumers prioritize transparency and sustainability over AI assistant performance?

**Tomohiro Ioku[1], Jaehyun Song[2] and Eiichiro Watamura[3]**

**Abstract**

As artificial intelligence (AI) becomes more integrated into society, concerns have arisen about unintended biases in AI-driven decision-making and the environmental impact of AI technology development. AI assistants such as Siri and Alexa, while helpful, can obscure decision-making and contribute to increased energy use and CO2 emissions. The present study explores whether consumers prioritize transparency and environmental sustainability over performance when choosing AI assistants with conjoint designs. Japanese participants were presented with different AI assistant profiles, varying in performance quality, transparency, cost, and environmental efficiency. The results revealed that Japanese participants prioritized transparency over performance when choosing AI assistants, but they prioritized performance over environmental sustainability. Moreover, future-oriented participants placed more importance on sustainability than those with a present orientation, while participants with an internal locus of control valued transparency more than those with an external locus of control. The findings of this study enhance our understanding of how consumers choose AI options and offer valuable guidance for creating AI systems and communication strategies that work effectively.

**Keywords**

Transparency, sustainability, conjoint analysis, social justice, locus of control, artificial intelligence

## Introduction

Artificial intelligence (AI) has become prevalent across in many sectors, such as transportation (Tham et al., 2022), healthcare (Stead, 2018), waste sorting (Aoki, 2020), esports games (Yokoi and Nakayachi, 2024), and legal systems (Mukai et al., 2023; Watamura et al., 2023). Artificial intelligence assistants such as Siri, Google Assistant, and Amazon Alexa are prominent devices performing various home tasks, including scheduling and information provision. However, using AI raises concerns about unintentional decision-making biases and serious environmental consequences (Floridi et al., 2018).

AI assistants' environmental and transparency issues are substantial yet often hidden. Their algorithms, which are not transparent to users, can obscure decision-making, influencing personal autonomy and producing a dependence on AI without sufficient understanding (Yeung, 2018). Moreover, AI technologies' energy and resource consumption contributes to increased energy use and $CO_2$ emissions (Dauvergne, 2021; Strubell et al., 2019), a concern often imperceptible to end-users. In response to these challenges, the European Union and the Ministry of Internal Affairs and Communications of Japan have implemented policies encouraging sustainable and transparent AI practices (European Commission, 2021; Ministry of Internal Affairs and Communications, 2022).

The present study investigates whether consumers prioritize transparency and environmental sustainability over performance when choosing an AI assistant, thereby addressing the gap in the existing research on consumer preferences in situations requiring trade-offs. Trade-off studies are critical in AI ethics, especially when weighing transparency against other principles, including

[1]Center for International Education and Exchange, Osaka University, Suita, Japan
[2]Faculty of Informatics, Kansai University, Takatsuki, Japan
[3]Graduate School of Human Sciences, Osaka University, Suita, Japan

**Corresponding author:**
Tomohiro Ioku, Center for International Education and Exchange, Osaka University, 1-1 Yamadaoka, Suita 565-0871, Japan.
Email: t_ioku@outlook.com

performance (Kieslich et al., 2022; König, Felfeli et al., 2022; König, Wurster et al., 2022). Enhancing transparency in AI means simplifying algorithms, potentially reducing their performance (Gunning et al., 2019; Köbis and Mossink, 2021; Nussberger et al., 2022); high-performance AI models are often complicated and less transparent. Therefore, our study aimed to contribute to trade-off studies in AI ethics by experimenting with a conjoint analysis design through a social-psychological approach.

## *Challenges of AI implementation: transparency and sustainability*

Extensive research in AI ethics covers various issues, such as fairness, discrimination, and accountability (Jobin et al., 2019); the present study focuses on two specific aspects: the lack of transparency in AI assistants and the rising energy consumption of AI, which presents environmental challenges. By focusing on these aspects, we aim to expand our understanding of their implications to contribute to the ethical implementation of AI technology.

Transparency in AI assistants is crucial, as it influences users' understanding and interaction with these technologies. Prior studies have highlighted the importance of transparency in algorithmic systems (Shin et al., 2022; Shin and Park, 2019). Shin and Park (2019) suggested that transparency shapes users' perceptions and interactions with these systems. Further, Shin et al. (2022) introduced "transparent fairness," a key concept for user interpretations of AIs' quality and credibility on media platforms, underscoring the role of user-centered transparency that combines technical clarity with user insights.

High transparency is essential for understanding AI technologies, as it facilitates users' awareness of decision-making processes and potential alternative outcomes (Edwards and Veale, 2017; Gunning et al., 2019; König, Wurster et al., 2022; Samek and Müller, 2019). Note that merely disclosing AI's general principles and methods is inadequate for users to gain a meaningful understanding (Felzmann et al., 2019). Users need detailed insights into how the system generates specific outputs, such as recommendations or decisions (Edwards and Veale, 2017; Malgieri and Comandé, 2017). This calls for a shift toward higher transparency via explainable AI, allowing users to understand alternative outcomes in varying conditions (König et al., 2023; Samek and Müller, 2019).

Effective AI engagement also requires a "justification approach" to transparency, which offers adequate, pertinent information for user acceptance without overloading them with technical details (de Fine Licht and de Fine Licht, 2020; Hilton, 1990; Miller, 2019). Miller (2019) emphasizes the importance of aligning the volume and type of information with user needs for meaningful interaction. Specifically, considering relevance, the amount of information, its epistemic value, and avoiding information overload is essential for tailoring explanations to users' specific needs (Hilton, 1990; Miller, 2019). Previous studies have explored how varying transparency levels in AI decision-making influence public perception and legitimacy (e.g., de Fine Licht and de Fine Licht, 2020). De Fine Licht and de Fine Licht (2020) advocate for a limited form of transparency that provides justifications for decisions, which may foster trust while preventing the public from being overwhelmed with technical details. Consequently, a justification approach is crucial for enhancing user engagement with AI, ensuring that transparency effectively bridges the gap between AI functionalities and user understanding.

On the environmental front, AI's impact is a serious concern due to its resource and energy demands. While AI can improve environmental sustainability by enhancing manufacturing efficiency through industrial robots—which reduces energy use and promotes sustainable technical innovations (Lee et al., 2022; Vinuesa et al., 2020)—its rapid growth is driving a steep increase in energy consumption (Dauvergne, 2021; Strubell et al., 2019). The Information and Communication Technology sector, including AI, accounts for approximately 2% of global $CO_2$ emissions, equal to the aviation industry (Strubell et al., 2019). Data centers, crucial for AI operations, consume 1% of the world's energy, with a notable recent surge (IEA, 2020).

The escalating global use of AI also raises concerns about worsening environmental issues (Cowls et al., 2023; König, Wurster et al., 2022). As AI becomes more prevalent, computing power demand could intensify environmental problems (Cowls et al., 2023). This is further complicated by the public perception of AI's environmental impact. Prior research on personal AI assistants revealed consumers' tendency to neglect environmental efficiency, specifically AI assistants' energy consumption (König, Wurster et al., 2022). This highlights the gap between public perception and the environmental necessity of prioritizing sustainability in AI choices.

Taken together, addressing AI's transparency and rising energy consumption issues is crucial for maintaining individual autonomy and promoting environmental sustainability. While these issues are the focus of our study, we acknowledge the importance of other ethical aspects of AI and stress the need for urgent attention and action in these areas.

## *Perceptions of AI transparency and sustainability: procedural and distributive justice framework*

The expectations of which features people perceive as more important can be derived through the lens of social justice theory (Colquitt and Rodell, 2015; Greenberg, 1990; Kordzadeh and Ghasemaghaei, 2022; Thibaut and

Walker, 1975; Tyler, 2000). Social justice theory has its roots in classic social-psychological theory, such as the work of Kurt Lewin, and explains how people's perceptions of justice shape their behavior and attitudes, including acceptance of decisions and systems within various social contexts (Gold, 1999; Tyler et al., 2015). Since Thibaut and Walker (1975) proposed a theory about the psychology of procedural preference, extensive research has focused on two dimensions of social justice theory: distributive and procedural justice (Hayashi, 2007; Ohnuma et al., 2022; Tyler, 2000).

Distributive justice refers to the perceived fairness of outcomes (e.g., performance quality, clinical diagnoses, or college admissions), which is based on principles of equality, need, or equity (Buijsman, 2023; Landers and Behrend, 2023; Lenzi et al., 2023; McFarlin and Sweeney, 1992; Tyler, 2000). People perceive a decision outcome to be distributive just when outcomes are expected to be distributed equally to all (equality principle), given to those who need them most (need principle), or proportional to inputs (equity principle). For example, applying the equity principle to AI systems, people may see the energy efficiency of an AI assistant as distributive just if its energy consumption and resource use are minimized and fairly distributed, ensuring no community, present or future, bears disproportionate energy costs or environmental impact.

Procedural justice refers to the perceived fairness of the process used to determine outcomes, involving transparency and consistency (Colquitt and Rodell, 2015; Grimmelikhuijsen, 2023; Tyler, 2000). It is independent of distributive justice, meaning the process can be seen as fair even if the outcome is unsatisfactory. People consider a decision process to be procedurally just when the criteria and methods are clearly explained (transparency) and uniformly applied across all cases (consistency). For example, AI assistants can be considered procedurally just by clearly and transparently explaining that restaurant suggestions are based on user ratings, reviews, and proximity.

A large body of research has investigated the importance of distributive and procedural justice (Besley, 2010; Earle and Siegrist, 2008; Krütli et al., 2012; Ohnuma et al., 2022; Sunshine and Tyler, 2003; Tyler and Caine, 1981). An experimental study shows that Swiss people generally trust public officials making decisions about radiation technologies if they believe the decision-making process is fair, irrespective of the specific decisions made (Earle and Siegrist, 2008). A survey study shows that Americans are more likely to accept decisions about nuclear technology if they receive an equal share of benefits or risks from nuclear power, and they are even more likely to accept such decisions when they believe the decision-making process is transparent (Besley, 2010). A case study in Japan, relevant to our context, shows that participants in programs for the development of sustainable policies who expect positive outcomes tend to accept these policies (Ohnuma et al., 2022). Furthermore, this acceptance is more robust when participants perceive the policy-making process as transparent, highlighting the importance of procedural justice in shaping public attitudes and behaviors in a Japanese context. These studies suggest that procedural justice primarily shapes our attitudes and behaviors rather than distributive justice; we call this the supremacy effect of procedural justice.

Recently, the importance of distributive and procedural justice has received attention in algorithmic decision-making (Acikgoz et al., 2020; Kieslich et al., 2022; Lee et al., 2019; Marcinkowski et al., 2020; Shin, 2020). Experimental studies reveal that when Americans perceive AI in the employee selection process as just, they are more likely to view the organization favorably and to pursue employment there, while being less likely to consider legal action against the organization (Acikgoz et al., 2020). This underscores the role of perceived procedural justice in algorithmic decision-making in employment contexts. Further, a survey study shows that American users of algorithmic personalized recommendations tend to trust algorithms that perceived as transparent (Shin, 2020). Another survey study finds that university students in Germany who view the results of an algorithmic admission system as equitable tend to support the system and positively rate the university's reputation (Marcinkowski et al., 2020). In contrast, those who perceive the system process as just tend to support the system and apply it to the university, demonstrating the importance of both distributive and procedural justice in educational settings. These studies highlight the importance of procedural and distributive justice in accepting AI technologies, yet few have directly compared their relative importance, with only a few exceptions.

More recently, there has been a notable work on algorithmic decision-making in public sector contexts: predictive policing and skin cancer risk prediction (König, Felfeli et al., 2022). The contexts involve predictive risk algorithms (for burglary or skin cancer) and indicate increased screening (more surveillance, more skin cancer screening) as the policy solution. This study examined how German people responded to trade-offs between the transparency of an algorithm and its performance. The results indicated a preference for performance over transparency, which seems contrary to the supremacy effect of procedural justice. This preference for performance can be understood as rational in immediate risk scenarios, such as healthcare, where accurate risk assessments are crucial because failing to detect skin cancer could have severe consequences. However, the preferences may differ in nonimmediate risk scenarios, such as when AI assists with entertainment, events, or product recommendations. In these contexts, transparency might be valued more highly than performance, as evidenced by the supremacy effect of procedural justice in other scenarios involving nonimmediate risk, such as policy development (Besley, 2010; Krütli et al., 2012;

Ohnuma et al., 2022; Tyler and Caine, 1981). Therefore, we hypothesize the following:

**H1.** Individuals perceive transparency as more important than performance quality and sustainability.

While some research divides justice into multiple dimensions—including distributive, procedural, informational, and interpersonal (Colquitt et al., 2001)—our study specifically examines procedural and distributive justice. This focus pertains to the research question of people's preferences when choosing an AI assistant. The tendency to favor transparency over performance and environmental sustainability corresponds with social justice theories, which often highlight procedural aspects over outcomes (Sunshine and Tyler, 2003; Tyler, 2000). Moreover, a few studies underscore the importance of procedural justice, particularly in transparency and control in AI design, which is crucial for users' perception of fairness (Lee et al., 2019; Marcinkowski et al., 2020). Similarly, distributive justice concerning AI outcomes, including performance and sustainability, can be equally crucial (though perhaps not as crucial as procedural justice) in user acceptance of AI systems, as noted earlier. Therefore, our study focuses on procedural and distributive justice, which is consistent with the research question on individuals' preferences in choosing AI assistants.

While theories such as the social construction of technology (Bijker, 1995; Dolata and Schwabe, 2023; Pinch and Bijker, 1984) and social practice theories (Shove and Pantzar, 2005; Shove and Walker, 2010) also explain the adoption and use of new technology, social justice theory is particularly relevant for understanding consumer preferences in AI. This is because it directly addresses ethical concerns in AI, unlike other theories focusing more on social and cultural influences (Bijker, 1995; Dolata and Schwabe, 2023; Shove and Walker, 2010). Additionally, social justice theory uniquely evaluates outcomes (distributive justice) and processes (procedural justice), which makes it crucial for understanding AI user preferences. For example, recent studies suggest its effectiveness in exploring how perceptions of fairness impact acceptance of technology (Lee et al., 2019; Starke et al., 2022). Thus, social justice theory can be a useful framework for investigating why consumers might prioritize transparency in AI systems.

## Perceptions of AI cost: pricing framework

In addition to distributive and procedural justice, pricing can be useful for understanding consumer perceptions around AI. It represents the compensation a customer offers for a product (Fan et al., 2022; Schindler, 2011). Two predominant models explain consumer reactions to pricing: the standard economic model and the zero-price model (Shampanier et al., 2007).

The standard economic model assumes rational consumer behavior, where decisions are made by weighing the cost against the benefits. For example, consumers are expected to prefer a $30 AI assistant subscription over its usual $50 rate due to perceived increased value from the $20 discount. In contrast, the zero-price model posits that free products disrupt rational decision-making because the absence of cost makes the product disproportionately appealing. For instance, even if a paid product offers more features, consumers might opt for a free version due to the allure of obtaining something without cost.

Shampanier et al. (2007) find that consumers overwhelmingly prefer products offered for free, suggesting that a zero-price enhances perceived value and demand, deviating from traditional economic predictions based on cost–benefit analysis. Subsequent studies across various products—such as food (Hossain and Saini, 2015; Palmeira and Srivastava, 2013), cosmetics (Spiegel et al., 2011), hotel packages (Zhang et al., 2023), and online services (Hüttel et al., 2018; Niemand et al., 2019)—have supported the zero-price model, reinforcing its validity in explaining consumer behavior divergent from purely economic rationality.

In AI technologies, users often prefer free or cost-effective options, as seen in free options used by software platforms (Niemand et al., 2019) and AI assistants (König, Felfeli, et al., 2022). This preference highlights the importance of considering not only transparency and environmental sustainability but also cost when examining how users weigh the features of AI technologies against their monetary costs.

## Individual differences in AI preferences

We also investigate different psychological profiles to better understand how individual differences influence user preferences for AI assistant features, such as transparency and environmental sustainability. Psychological profiles can impact how users perceive and interact with AI technologies, and identifying these differences can help tailor AI systems to meet diverse user needs effectively. As detailed below, we chose locus of control and future orientation due to their relevance: individuals with an internal locus of control value transparency to understand and influence systems, while future-oriented individuals prioritize long-term benefits, including environmental sustainability.

*Locus of control.* The concept of locus of control, rooted in Rotter's social learning theory (Galvin et al., 2018; Rotter, 1954, 1966; Wang et al., 2010), provides a framework for understanding individual differences in AI technology perception. Individuals with an internal locus of control believe they can influence events and outcomes in their lives (Spector, 1982). Such internals exhibit confidence and determination, attributing their outcomes to

personal efforts (Allen et al., 2005; Rotter, 1966). This contrasts with externals, who perceive outcomes as products of external forces (Ng et al., 2006; Spector, 1982).

Because of their sense of control and self-efficacy, internals may perceive transparency in AI assistants as more crucial than externals. This preference toward transparency corresponds with the internals' desire for mastery and control, as evidenced in their approach to technology (Ryff, 1989; Zimmerman and Rappaport, 1988). Transparent AI assistants offer visibility into the technology's workings, aligning with the internals' preference for understanding and influencing the system they interact with.

Supporting this notion, previous studies show that user control, a preference for internals, may increase satisfaction, possibly due to reduced perceived demands and frustrations in system interactions (Dietvorst et al., 2018). This is in line with research suggesting that high work demands can result in mental health issues, a concern mitigated by user control (McClure, 2018). Additionally, prior beliefs, such as internal locus of control, influence perceptions of algorithm-aided systems in various contexts (Prahl and Van Swol, 2017; Promberger and Baron, 2006). This further supports the idea that the internals' perception of control extends to interaction with technology, making transparency a more valued feature. Therefore, our second hypothesis is as follows:

**H2.** Individuals with an internal locus of control perceive transparency as more important than those with an external locus of control.

*Future orientation.* Previous research on future orientation can shed light on which individuals are more likely to consider certain features important. Future orientation refers to how individuals think about the long-term outcomes of their present actions and how these possible outcomes influence their present actions (Corral-Verdugo and Pinheiro, 2006; Strathman et al., 1994). The concept of future orientation originated from the seminal work of Strathman et al. (1994) and is closely tied to the construal-level theory (Engle-Friedman et al., 2022; Gu et al., 2020).

According to construal-level theory, individual differences exist in the extent to which a person feels separate in the occurrence, time, geographical space, or social connection from an event, which is called psychological distance (Trope and Liberman, 2010). Psychological distance to an event increases the difficulty in conceptualizing the event (Jones et al., 2017; Wakslak and Trope, 2009), suppressing an affective reaction to the event and reducing the likelihood of preparatory action (Engle-Friedman et al., 2022). As construal-level theory assumes, individuals exhibit varying degrees of consideration for long-term outcomes when making choices (Strathman et al., 1994). Some naturally prioritize future benefits, even if there are immediate costs. These individuals with future orientation are willing to sacrifice immediate gratification or convenience to achieve more desirable future states.

Given construal-level theory, future-oriented individuals may prioritize environmental sustainability when choosing an AI assistant; they tend to engage in various pro-environmental behaviors (Corral-Verdugo and Pinheiro, 2006; Enzler et al., 2019; Joireman et al., 2004). For example, consider a person with a future orientation who wants to incorporate more sustainable transportation practices into their daily life. They search for an AI assistant that encourages eco-friendly commuting. Their future orientation drives them to actively engage in pro-environmental behaviors and use the AI assistant to make sustainable choices. A recent study reveals that future-oriented individuals show a preference for regulatory measures regarding AI's ecological sustainability (König et al., 2023). König et al. (2023) found a correlation between future orientation and support for both strict (hard) and flexible (soft) regulatory frameworks for AI using a German sample. This further highlights how future orientation not only influences personal choices but also extends to preferences for broader societal and regulatory approaches to AI and environmental sustainability.

However, there might be doubt about whether future orientation motivates people to choose environmentally friendly AI assistants. Due to the unpredictable nature of the future (Brügger et al., 2015), people are often reluctant to invest in it. Moreover, there is a tendency known as time discounting, where the perceived value of future events diminishes over time (Ballard and Knutson, 2009; Engle-Friedman et al., 2022; Lünich et al., 2021). Put differently, we value positive outcomes more if they occur now rather than later, and the negative impact of future events seems to diminish the further away they are. This effect of time discounting grows stronger with increasing time intervals, influencing immediate decisions (Frederick et al., 2002). In digital media and online privacy, this translates to users favoring short-term advantages over potential long-term risks, showing a preference for immediate rewards and less concern for future consequences, a trait common in those with a present-focused orientation (Lünich et al., 2021). Thus, it is essential to examine whether future orientation motivates environmentally friendly choices in AI assistants, as time discounting might induce a preference for immediate benefits over long-term sustainability in digital and AI-related decision-making. Our third hypothesis is as follows:

**H3.** Future-oriented individuals perceive environmental sustainability as more important than those with present orientation.

## The present study

We conducted a conjoint experiment to gather accurate information about people's preferences for AI assistants.

This method allowed us to understand the trade-offs made when individuals consider the different features of an AI assistant based on their choices across various assistants. In line with previous studies (König, Wurster et al., 2022), we examined four features of AI assistants in our research: (1) user satisfaction as an indicator of performance quality, (2) environmental sustainability, represented by the energy consumption of data centers when using the AI assistant, (3) transparency level, and (4) costs. Our study involved nearly 1000 respondents from a Japanese online panel, which provided a solid foundation for reliable analysis and allowed us to conduct subgroup evaluations.

Our study shares similarities with that of König, Felfeli et al. (2022), who also explored whether people prioritize transparency over performance. However, while König, Felfeli et al. (2022) focused on public sector algorithms in policing and healthcare contexts, our study extends this inquiry to the context of personal AI assistants used in everyday life. As König, Felfeli et al. (2022) noted, individuals may have varying evaluations in different decision-making contexts, particularly when they have experienced negative impacts from algorithms (Schiff et al., 2022). Furthermore, while previous studies on social justice theory suggest a preference for transparency over performance, this notion appears to contrast with the findings of König, Felfeli et al. (2022). Such discrepancies require more research on the preference for transparency versus performance in different contexts.

Further, this study extends previous studies by addressing the gap in the fairness of algorithms. Previous studies predominantly used a variety of methods mainly conducted in Western democracies, such as the U.S. and Germany (König, Felfeli et al., 2022; König, Wurster et al., 2022; Marcinkowski et al., 2020; Shin, 2020; Starke et al., 2022). The issues with generalizing findings from primarily WEIRD (White, Educated, Industrialized, Rich, Democratic) samples are well-recognized (Henrich et al., 2010). As fairness in decision-making varies significantly across different sociocultural contexts (Henrich et al., 2010), perceptions of AI fairness vary widely across different countries (Kelley et al., 2021). Accordingly, including countries like Japan, one of the leading East Asian countries in responsible AI development and governance (Habuka, 2023) could contribute to the existing literature.

More importantly, the present study not only addresses gaps in the fairness of algorithms but also advances the field by uniquely examining how individual differences, such as locus of control and future orientation, shape preferences for AI assistant features. While some research has explored how much people value specific features of AI systems, such as fairness (König, Felfeli et al., 2022; Shin, 2020; Shin et al., 2022), our work integrates existing theories on locus of control (Rotter, 1954, 1966) and future orientation (Strathman et al., 1994) to offer new insights. This integration provides a deeper understanding of how preferences for transparency and environmental sustainability in AI assistants depend on these individual characteristics.

## Method

This study was approved by the Research Ethics Committee of the Graduate School of Human Sciences at Osaka University. All data, materials, and analysis code can be found in the Open Science Framework: https://osf.io/hf7za/?view_only=0f5b30e1a7c5494a8e74e080232f39a2.

### *Participants*

Participants were recruited via Yahoo! Crowd Sourcing from April 18, 2023, to April 25, 2023. The criteria for inclusion were being a Japanese resident and 18 or older. We initially received 954 responses. To ensure participants understood this study, we administered three quizzes on AI assistants, transparency, and energy consumption, each with two answer choices and explanatory text. Following the exclusion of incorrect quiz responses, our final sample comprised 833 participants: 539 males, 286 females, and 8 identifying as other. The average age was 48.51 years ($SD = 11.37$).

### *Conjoint analysis*

To evaluate how much importance individuals place on AI assistants' transparency and environmental sustainability, we conducted a conjoint experiment within a survey. Participants chose between various AI assistant profiles with multiple features rather than answering direct questions about specific features. This method allowed us to deduce the utility of each option and feature from the choices made, helping to mitigate social desirability bias (Hainmueller et al., 2014) and reflecting real-world scenarios where trade-offs are common. Participants were presented with a randomized set of nine choice tasks. In each task, they were instructed to choose one AI assistant from a set of two. These options varied in terms of four features: performance quality—measured by the percentage of users satisfied with the AI assistant's recommendations; transparency of the AI assistant; monthly costs; and environmental efficiency, captured by external energy consumption. Regarding performance quality, when evaluating the AI assistants, we informed participants that user satisfaction with the recommendations provided is a crucial indicator. We set the lowest level of satisfaction at 89% (just below 90%) to ensure a noticeable distinction from the second level, set at 94%. The highest level was 99% satisfaction. In terms of transparency, our feature levels included no transparency, low transparency through basic functionality disclosure, and high transparency through explainable AI. The highest level provides users with additional information about the reasoning behind any recommendations

made, empowering users to make informed decisions. Considering the cost aspect of AI assistants, we considered the common expectation that mobile AI assistants are often free; hence, the first level represents no costs. We presumed that higher monthly fees, such as 300 Yen and 600 Yen for the second and third levels, respectively, would be perceived as unfavorable. Although these amounts may not be substantial in absolute terms, they are not minor compared to typical mobile contract expenses. Lastly, for the energy efficiency of AI assistants, we based the feature levels on the estimated $CO_2$ footprint of a search machine query. Specifically, we expressed the $CO_2$ emissions in terms of the hours an energy-saving lamp can operate before emitting the same amount of $CO_2$ as 10 AI assistant queries. Feature levels represented the ability to run an energy-saving lamp for one, three, or five hours in exchange for 10 AI assistant queries.

### Measures

For subgroup analysis, we measured two constructs: locus of control, adapted from Levenson's scale (Levenson, 1973), and future orientation, adapted from Joireman et al.'s scale (Joireman et al., 2012). Locus of control was measured using items assessing the belief in internal (seven items) versus external factors (seven items) and rated on a seven-point Likert scale ranging from 1 (*Strongly disagree*) to 7 (*Strongly agree*). Sample items included "My life is determined by my own actions" and "My life is chiefly controlled by powerful others." Items for external factors were reversed. This scale showed good reliability, $\alpha = 0.77$. We created subgroups of internal and external locus of control using a mean-split method.

Future orientation was measured through participants' tendency toward present or future outlooks, using a similar seven-point scale ranging from 1 (*Not at all like you*) to 7 (*Very much like you*). Items included "I am willing to sacrifice my immediate happiness or wellbeing in order to achieve future outcomes" and "I generally ignore warnings about possible future problems because I think the problems will be resolved before they reach crisis level" Items for present orientations were reversed. This scale also demonstrated satisfactory consistency, $\alpha = 0.83$. Using a mean-split technique, we formed subgroups distinguishing between present and future orientation.

### Analyses

We processed our Qualtrics conjoint experiment data using the "*q2c*" package in R (Song, 2022), which processes data fast and corrects column misalignment, issues present in the "*read.qualtrics*" function in the "*cjoint*" package (Hainmueller et al., 2014). Our analysis utilized the "*cregg*" package to compute three essential metrics (Leeper et al., 2020): marginal means (MMs), average marginal component effects (AMCEs), and MM differences (MM diffs). AMCEs capture the effect of altering one attribute while holding others constant. MMs reflect the degree of preference for choices with a particular attribute level, ignoring all other attributes (Leeper et al., 2020). AMCE focuses on the relative effect of changing a feature compared to a baseline, useful for causal inference, while MM focuses on the absolute level of preference toward different feature levels, which makes it more straightforward for descriptive analysis. For instance, if the transparency attribute has an MM of 52.5% for high transparency and 48.5% for low transparency in AI assistants, this shows a preference of 52.5% for high transparency and 48.5% for low transparency. The AMCE for transparency would be 0.04 (52.5–48.5), suggesting that AI assistants with high transparency are 4% more likely to be chosen than those with low transparency. Finally, MM diffs measure the variances in MMs across subgroups, indicating the extent and significance of these differences. This is relevant for assessing whether preferences for AI features significantly vary among subgroups, such as those with an internal locus of control or future orientation.

## Results

### Analysis of the full sample

Figures 1 and 2, along with Tables 1 and 2, show the results from the complete data. The vertical axis represents various attributes and their levels. In contrast, compared with the base attribute levels, the horizontal axis shows AMCEs and MMs for each attribute on the likelihood of choosing an AI assistant. These base levels are defined as 89% satisfaction, no transparency, free cost, and the energy consumption of 10 queries equivalent to one hour of an energy-saving lightbulb's usage.

First, increasing user satisfaction from 89% to 94% (AMCE = 0.09) and then to 99% (AMCE = 0.15) significantly raised the probability of choosing an AI assistant. Participants chose AI assistants with 94% and 99% satisfaction at rates of 50.3% and 57.1%, respectively, while those with 89% satisfaction were chosen at a rate of 42.5%. Second, the probability of choice decreased when the energy consumption for 10 queries was equivalent to three or five hours of energy-saving lamp usage (AMCE = −0.03 and −0.07, respectively) as opposed to one hour. Participants chose AI assistants with energy usage equating to one hour and three hours of lamp usage at rates of 53% and 50.7%, while those with five-hour usage were chosen 46.5% of the time. Third, the probability of choosing an AI assistant increased with both low and high transparency (AMCE = 0.05 and 0.22) compared to no transparency. Participants chose high-transparency AI assistants 63% of the time, compared to 40.5% and 46.6% for those with no or low transparency, supporting Hypothesis 1. Finally, switching from no cost to a monthly fee of 300 and

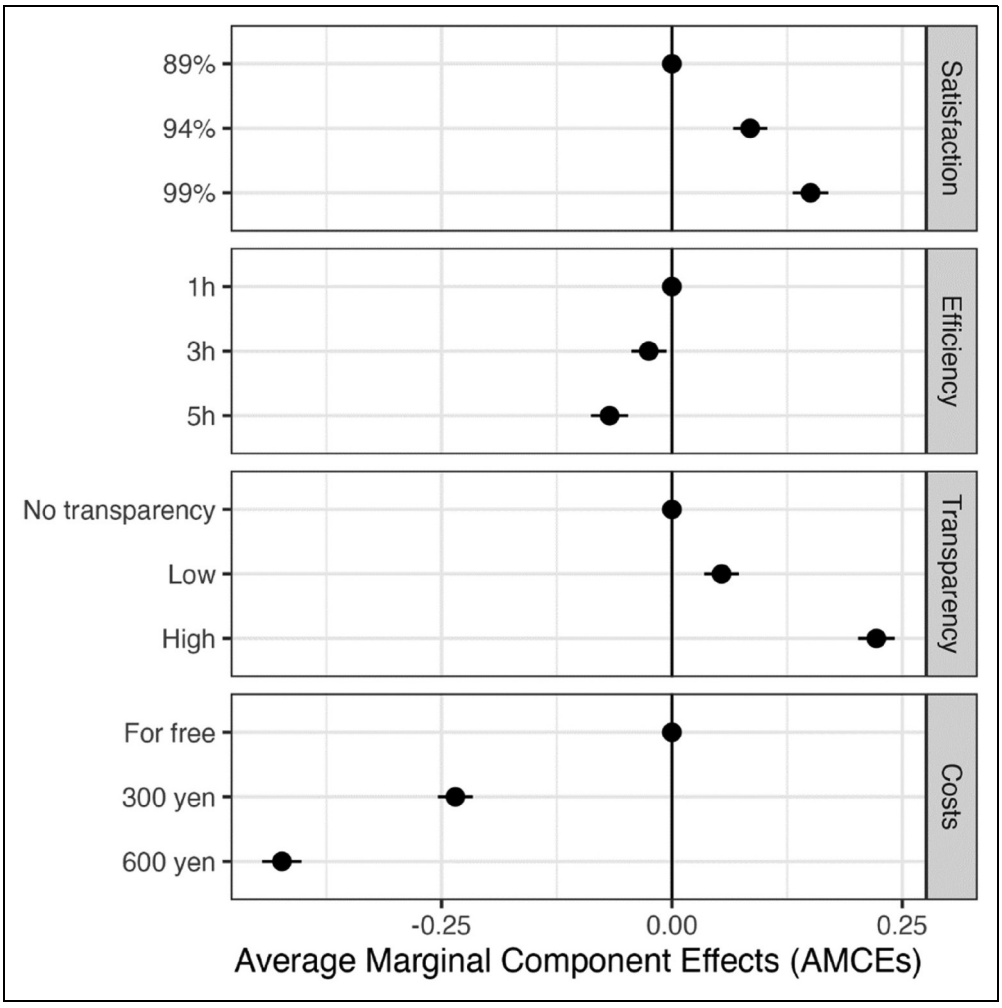


**Figure 1.** AMCEs on AI selections with a pooled sample.

600 Yen decreased the probability of choice (AMCE = −0.24 and −0.42). Participants chose free AI assistants 71.8% of the time, while they chose those priced at 300 Yen and 600 Yen at rates of 48.5% and 29.6%, respectively.

### *Subgroup analysis*

Figures 3 and 4 and Table 3 present the results from both internal and external groups. The internal group showed greater concern for transparency than the external group. Unlike other features, there was a significant difference in the effect of transparency ($p < .05$). For the internal group, high transparency had the conditional MM of 64.8% versus 61.8% for the external group, endorsing Hypothesis 2. This suggests that those with an internal locus of control value transparency more than those with an external locus.

Next, Figures 5 and 6, alongside Table 4, present the results for groups with different time orientations—present and future. The future-oriented group placed a higher value on AI's environmental sustainability more than the present-oriented group. We observed a significant difference in the effect of energy efficiency ($p < .05$). In the future-oriented group, the conditional MM for five-hour energy consumption was 44.6%, compared to 47.5% in the present-oriented group. This supports Hypothesis 3, showing that future-oriented people prioritize environmental sustainability more than present-oriented people. Interestingly, the future-oriented group also valued transparent AI more, which was an unexpected finding.

## Discussion

Our findings offer valuable insights into the importance individuals attach to AI assistants' transparency and environmental sustainability. The results of our analysis demonstrate that people prioritize transparency in AI assistants over environmental sustainability and performance. This preference for transparency is particularly evident among individuals with an internal locus of control or future orientation. Although the environmental sustainability of AI

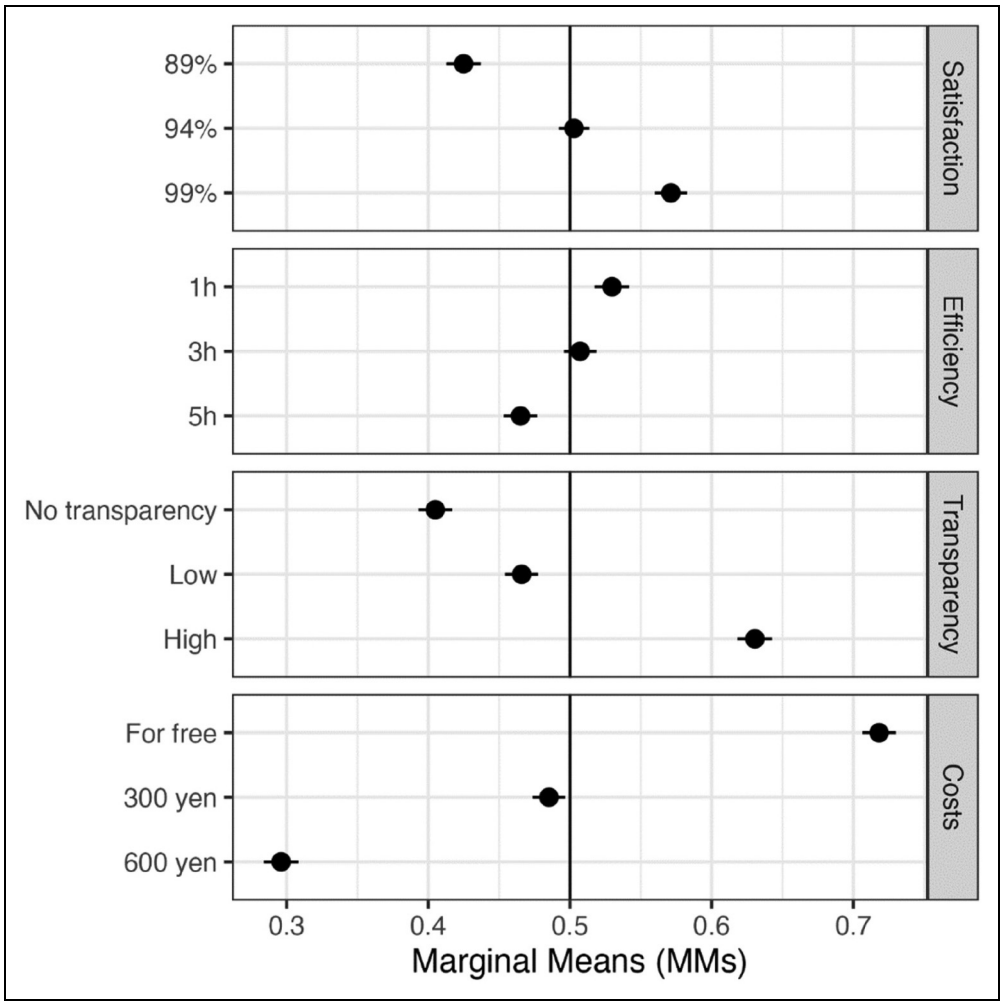


**Figure 2.** MMs on AI selections with a pooled sample.

**Table 1.** AMCEs on AI selections with a pooled sample.

| | | | | 95% CI | |
|---|---|---|---|---|---|
| | AMCE | SE | *p* | Lower | Upper |
| **Satisfaction** | | | | | |
| 89% (Baseline) | | | | | |
| 94% | 0.085 | 0.009 | <.000 | 0.066 | 0.103 |
| 99% | 0.150 | 0.010 | <.000 | 0.131 | 0.170 |
| **Efficiency** | | | | | |
| 1 h (Baseline) | | | | | |
| 3h | −0.025 | 0.010 | .009 | −0.044 | −0.006 |
| 5h | −0.068 | 0.010 | <.000 | −0.088 | −0.048 |
| **Transparency** | | | | | |
| No transparency (Baseline) | | | | | |
| Low | 0.054 | 0.010 | <.000 | 0.035 | 0.073 |
| High | 0.222 | 0.010 | <.000 | 0.202 | 0.242 |
| **Costs** | | | | | |
| For free (Baseline) | | | | | |
| 600 yen | −0.423 | 0.011 | <.000 | −0.445 | −0.402 |
| 300 yen | −0.235 | 0.010 | <.000 | −0.254 | −0.216 |

**Table 2.** MMs on AI selections with a pooled sample.

| | | | | 95% CI | |
|---|---|---|---|---|---|
| | MM | SE | *p* | Lower | Upper |
| **Satisfaction** | | | | | |
| 89% | 0.425 | 0.006 | <.000 | 0.413 | 0.437 |
| 94% | 0.503 | 0.006 | <.000 | 0.492 | 0.514 |
| 99% | 0.571 | 0.006 | <.000 | 0.560 | 0.583 |
| **Efficiency** | | | | | |
| 1h | 0.530 | 0.006 | <.000 | 0.517 | 0.542 |
| 3h | 0.507 | 0.006 | <.000 | 0.496 | 0.519 |
| 5h | 0.465 | 0.006 | <.000 | 0.453 | 0.477 |
| **Transparency** | | | | | |
| No transparency | 0.405 | 0.006 | <.000 | 0.393 | 0.417 |
| Low | 0.466 | 0.006 | <.000 | 0.454 | 0.478 |
| High | 0.630 | 0.006 | <.000 | 0.618 | 0.643 |
| **Costs** | | | | | |
| For free | 0.718 | 0.006 | <.000 | 0.706 | 0.730 |
| 300 yen | 0.485 | 0.006 | <.000 | 0.474 | 0.497 |
| 600 yen | 0.296 | 0.006 | <.000 | 0.284 | 0.308 |

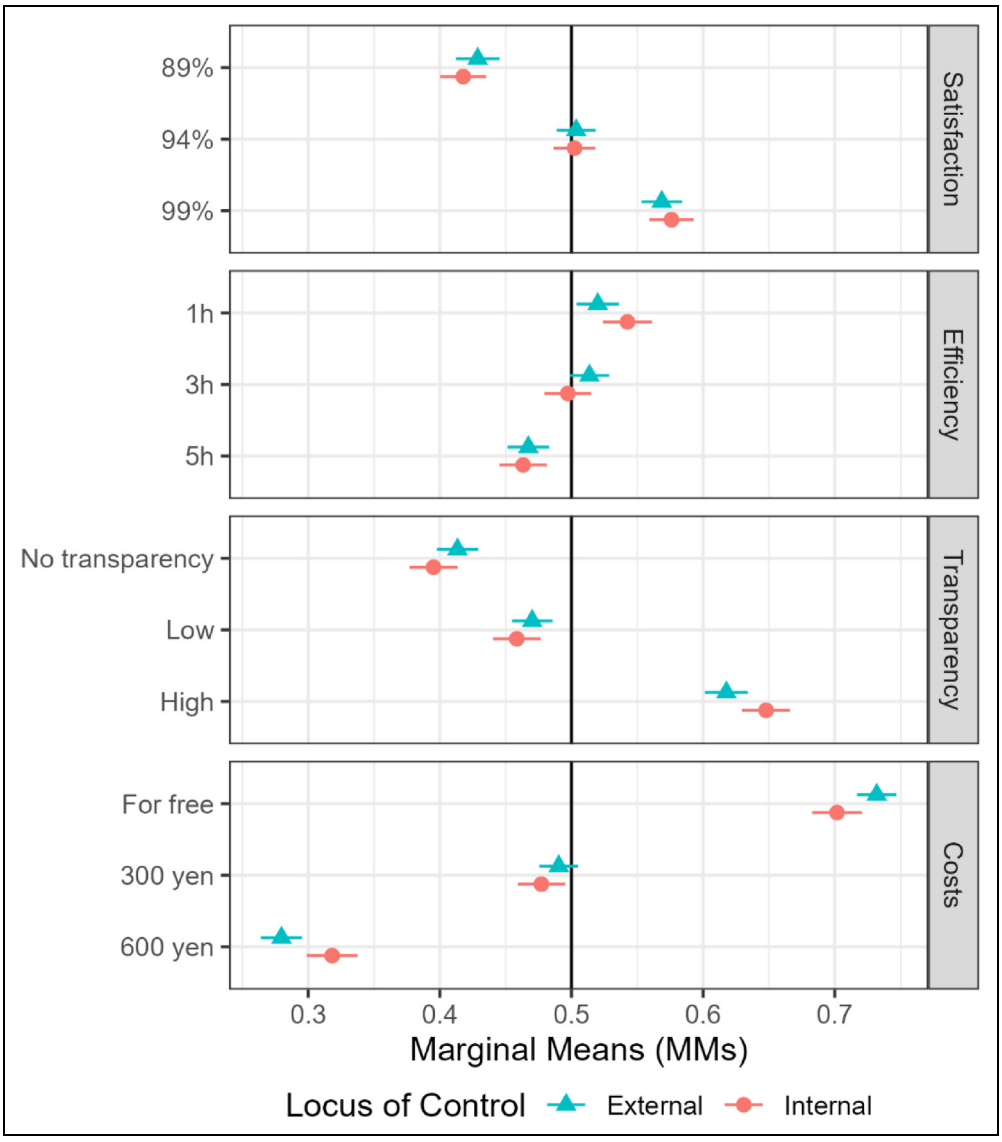


**Figure 3.** Conditional MMs on AI selections by locus of control.

assistants is of lesser concern to the Japanese population, it remains important, particularly among those with a future orientation. Our findings are consistent with previous studies that highlight the preference for transparency in AI assistants over performance quality and environmental sustainability (König, Wurster et al., 2022; Shin, 2021; Shin and Park, 2019). Understanding these preferences is crucial for developing and implementing AI systems that meet user expectations and reflect their values.

### *Theoretical implications*

Our results offer three major implications for algorithms, AI, and automation technologies theories. First, this study expands on the existing research about social justice studies (Starke et al., 2022; Tyler, 2000). Previous studies on social justice theory suggest the supremacy effect of procedural justice over distributive justice (Sunshine and Tyler, 2003; Tyler and Caine, 1981). An emerging body of research on the fairness of AI also supports the supremacy effect of procedural justice (König, Wurster et al., 2022; Marcinkowski et al., 2020; Shin, 2020). However, those studies were conducted in Western democracies. Our study extends these findings, revealing that transparency shaped Japanese preferences for AI assistants. This preference was more substantial than that for performance satisfaction and energy efficiency, which is consistent with Hypothesis 1. Therefore, our study supports the supremacy effect of procedural justice on preferences for AI technologies in a non-Western country.

Second, this study enhances the theoretical understanding of locus of control in AI technology perception,

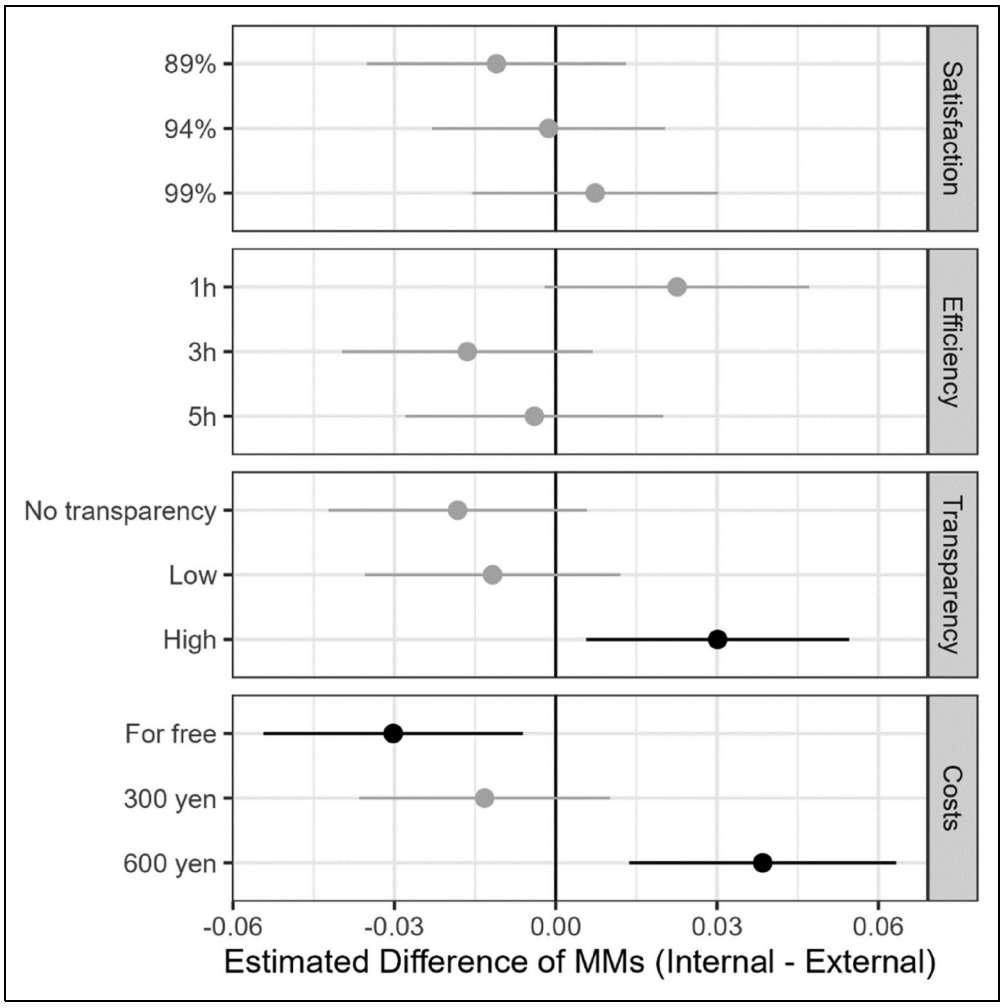


**Figure 4.** Differences in MMs on AI selections by locus of control.

drawing on the foundational work of Rotter (1954, 1966). We found that individuals with an internal locus of control value transparency in AI assistants. This preference reflects their inclination for mastery and control (Ryff, 1989; Zimmerman and Rappaport, 1988), in contrast to those with an external locus of control, who attribute outcomes to external factors (Ng et al., 2006; Spector, 1982). Our subgroup analysis reinforces this, showing that the internal group had a higher conditional MM for high transparency (64.8%) than the external group (61.8%), thus supporting Hypothesis 2. This underscores the notion that internals prioritize transparency more than externals in AI systems. Our findings are in line with the work of Dietvorst et al. (2018), suggesting that increased control correlates with greater satisfaction, which might lessen perceived demands and frustrations in AI interactions. Our study extends the discourse on locus of control, providing evidence for its role in shaping perceptions and interactions with technology.

Third, our study underscores the importance of future orientation in AI preferences, particularly regarding environmental sustainability. Building on previous research on future orientation (Corral-Verdugo and Pinheiro, 2006; Enzler et al., 2019; Joireman et al., 2004), consistent with construal-level theory (Trope and Liberman, 2010), our results reveal a clear preference for sustainable AI among future-oriented individuals. Our subgroup analysis shows that this group values AI's environmental sustainability more than the present-oriented group, as evidenced by the significant difference in responses to energy efficiency. The future-oriented group's conditional MM for AI systems with higher energy consumption was lower (44.6%) than that of the present-oriented group (47.8%), supporting Hypothesis 3. This preference for environmental sustainability is consistent with König et al. (2023), which highlights a preference for environmentally conscious regulatory measures in AI among future-oriented individuals, despite potential challenges posed by time discounting (Ballard and Knutson, 2009; Engle-Friedman et al., 2022). This suggests a role of future orientation in AI technology preferences and opens new avenues for further exploration of the psychological mechanisms behind these preferences.

**Table 3.** Conditional MMs on AI selections by locus of control.

| | | | | 95% CI | |
|---|---|---|---|---|---|
| | MM | SE | *p* | Lower | Upper |
| **External group:** | | | | | |
| **Satisfaction** | | | | | |
| 89% | 0.429 | 0.008 | <.000 | 0.412 | 0.445 |
| 94% | 0.504 | 0.008 | <.000 | 0.489 | 0.518 |
| 99% | 0.569 | 0.008 | <.000 | 0.553 | 0.584 |
| **Efficiency** | | | | | |
| 1h | 0.520 | 0.008 | <.000 | 0.504 | 0.536 |
| 3h | 0.514 | 0.008 | <.000 | 0.499 | 0.529 |
| 5h | 0.467 | 0.008 | <.000 | 0.451 | 0.483 |
| **Transparency** | | | | | |
| No transparency | 0.413 | 0.008 | <.000 | 0.398 | 0.429 |
| Low | 0.470 | 0.008 | <.000 | 0.455 | 0.485 |
| High | 0.618 | 0.008 | <.000 | 0.601 | 0.634 |
| **Costs** | | | | | |
| For free | 0.732 | 0.008 | <.000 | 0.717 | 0.747 |
| 300 yen | 0.490 | 0.008 | <.000 | 0.476 | 0.505 |
| 600 yen | 0.280 | 0.008 | <.000 | 0.264 | 0.295 |
| **Internal group:** | | | | | |
| **Satisfaction** | | | | | |
| 89% | 0.418 | 0.009 | <.000 | 0.400 | 0.435 |
| 94% | 0.502 | 0.008 | <.000 | 0.486 | 0.518 |
| 99% | 0.576 | 0.009 | <.000 | 0.559 | 0.593 |
| **Efficiency** | | | | | |
| 1h | 0.543 | 0.010 | <.000 | 0.524 | 0.561 |
| 3h | 0.497 | 0.009 | <.000 | 0.479 | 0.515 |
| 5h | 0.463 | 0.009 | <.000 | 0.445 | 0.481 |
| **Transparency** | | | | | |
| No transparency | 0.395 | 0.009 | <.000 | 0.377 | 0.413 |
| Low | 0.458 | 0.009 | <.000 | 0.440 | 0.477 |
| High | 0.648 | 0.009 | <.000 | 0.630 | 0.666 |
| **Costs** | | | | | |
| For free | 0.702 | 0.010 | <.000 | 0.683 | 0.721 |
| 300 yen | 0.477 | 0.009 | <.000 | 0.459 | 0.495 |
| 600 yen | 0.318 | 0.010 | <.000 | 0.299 | 0.338 |

## *Practical implications*

Our findings highlight the trade-offs when choosing AI assistants. Practical implications can be derived from our research, in that clear information about the trade-offs involved in choosing AI assistants should be provided. Developers should communicate the advantages and disadvantages of different features, such as performance quality, transparency, environmental sustainability, and cost, to help users make well-informed decisions, based on their individual preferences. By presenting the trade-offs, users can evaluate the relative importance of each feature and make choices that match their preferences.

Further, the results of our study can inform the development of more personalized AI systems. Acknowledging variations in how users perceive AI, such as their locus of control, can lead to more customized AI experiences. For instance, AI systems offering greater transparency and control might appeal more to those with an internal locus of control. This personalization in AI assistants could enhance user satisfaction and engagement by catering to individual preferences.

Furthermore, our findings have implications for encouraging pro-environmental behavior through AI. This study suggests that individuals with a future orientation are likely to engage in and value environmentally sustainable AI assistants. Artificial intelligence systems can be leveraged to promote and encourage sustainable practices, particularly among future-oriented individuals. Artificial intelligence assistants can empower users to make environmentally conscious choices and contribute to a sustainable future by providing personalized recommendations, real-time feedback, and adaptive learning algorithms. Further research can explore the effectiveness of different AI-driven interventions in promoting pro-environmental behaviors.

Finally, the present study also brings to light a crucial aspect often overlooked in discussions on AI ethics: the impact of cost on user choices. We reveal that while transparency and environmental sustainability are essential factors, cost is the most decisive feature for users when choosing AI assistants. No matter how transparent or eco-friendly AI algorithms may be, their adoption is heavily restricted when users face monetary costs. Consistent with König, Wurster et al. (2022), this suggests that users prefer an AI assistant with transparency or environmental sustainability only when the technology is free of charge. The probability of technology adoption plummets when a cost is associated with it, regardless of its ethical attributes. This poses a serious limitation for the general pursuit of AI ethics, implying that depending solely on consumer choice to ensure ethical AI practices is insufficient. These findings suggest that policymakers and AI developers must consider the economic aspects of AI adoption alongside ethical features. To ensure the implementation of AI ethics in a way that is both accessible and appealing to the public, strategies such as subsidizing ethical AI technologies or providing incentives for choosing ethical options might be necessary.

## *Limitations and future directions*

Focusing on user preferences for AI assistants, our study encounters several limitations that call for careful interpretation of the results and suggests areas for future research. First, the participant pool from Yahoo! Crowd Sourcing might not adequately reflect the wider Japanese population. Variations in age, gender, and socioeconomic factors could have skewed our findings, reducing their applicability to the public. We expect future studies with a representative sample to test the robustness of this study's findings.

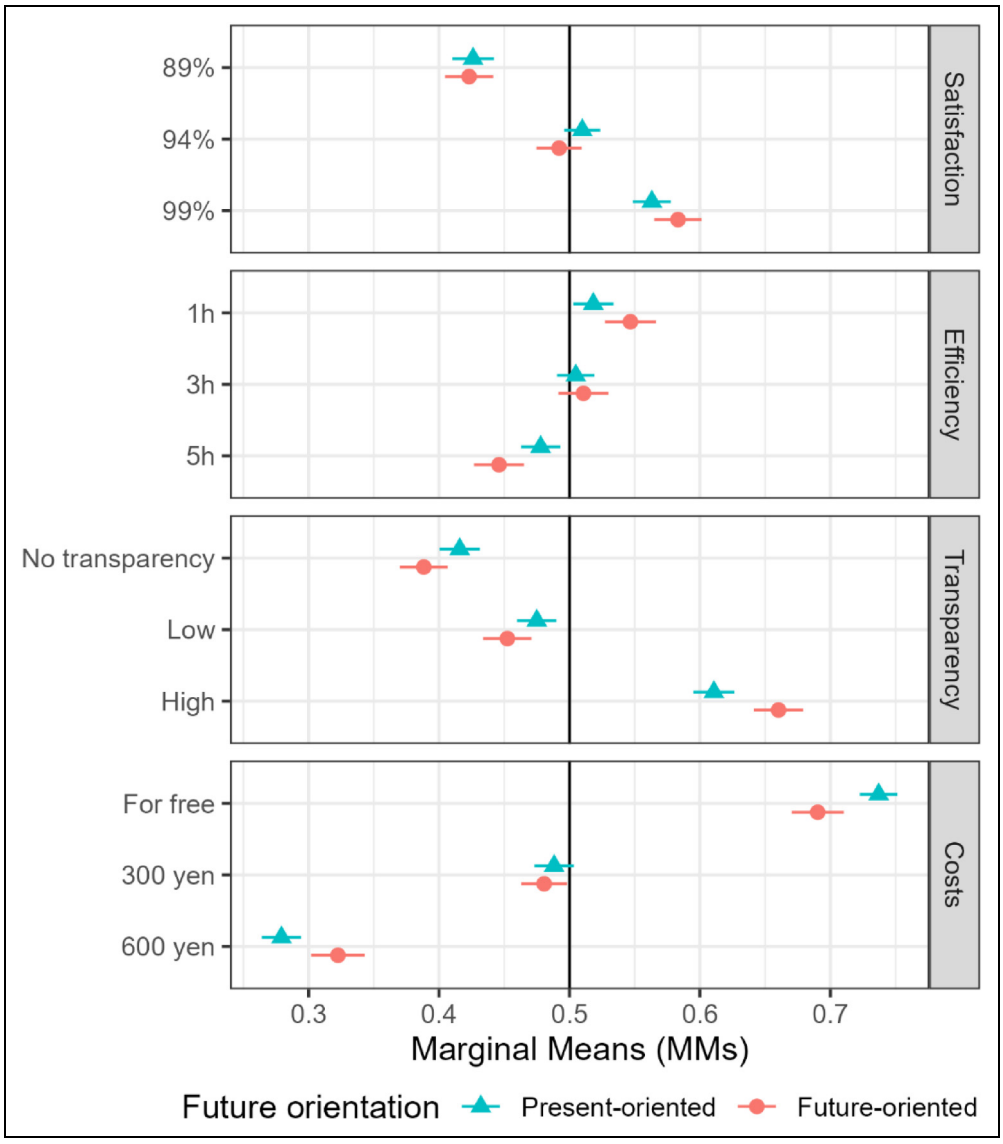


**Figure 5.** Conditional MMs on AI selections by future orientation.

Second, the chosen satisfaction levels of the AI assistants (89%, 94%, and 99%) were relatively high. The lowest level still implies that a vast majority, nearly 9 out of 10 users, are satisfied with the AI assistant. This restricted range could limit the detection of nuanced differences in user preferences. The difference between 94% and 99% satisfaction might not be meaningful to users, similar to high-rated hotel reviews, where the difference between an 89% and 94% rating is minimal. Future studies should use a broader spectrum of satisfaction levels to accurately capture user preferences and the effect of performance satisfaction on AI assistant choice.

Third, the method to quantify energy efficiency, based on the runtime of an energy-saving lamp, may have been too abstract or insignificant for participants to associate with AI's environmental impact. Given the low-energy consumption of such lamps, the small differences in operation time (one, three, or five hours) might seem negligible to the user. This could explain the minor effect of energy efficiency in our study. Future research should employ more relatable and meaningful metrics to better represent the environmental aspects of AI technologies, ensuring they associate more with users' daily experiences.

Finally, our study unexpectedly found that individuals with a future orientation gave higher importance to transparency in AI assistants. This was not predicted, and the rationale behind this result remains to be determined. While this adds a new perspective to the literature on AI preferences, we need an explanatory framework to

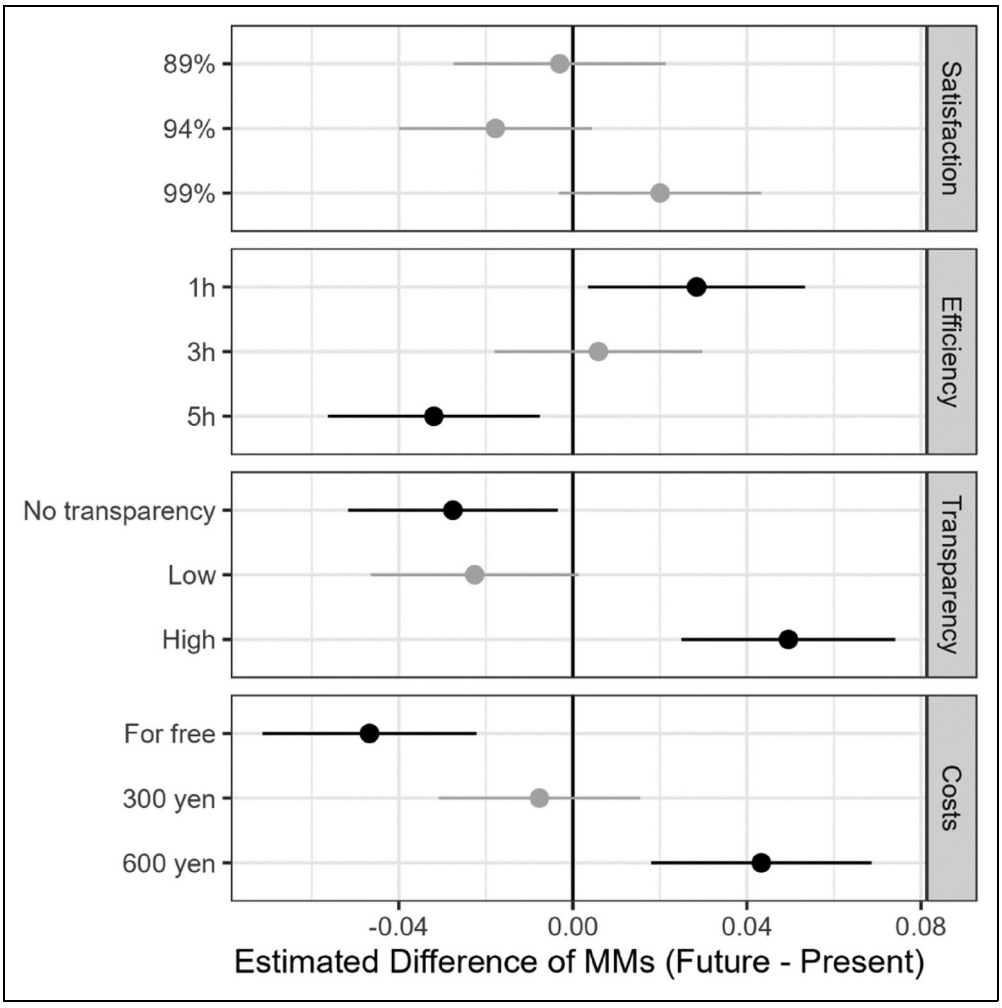


**Figure 6.** Differences in MMs on AI selections by future orientation

expand our understanding of the underlying factors. Future researchers should explore the psychological and behavioral reasons linking future orientation with a preference for transparency in AI technologies.

## Conclusion

The evolution of AI, particularly AI assistants, has brought unprecedented convenience and efficiency to our daily lives. However, the rapid adoption and integration of these technologies have also stirred concerns related to transparency and environmental sustainability. This study aimed to shed light on these critical issues by exploring how individuals value transparency and environmental sustainability in relation to other features when choosing an AI assistant. Our findings provide valuable insights into these trade-offs. We found that while individuals consider transparency and environmental sustainability when making their decisions, these factors are secondary compared to the overriding concern of cost. This means that, although these attributes are noteworthy, they are not the key decisive factors in choosing AI assistants. The results of this study underscore the importance of these attributes and suggest that they should be considered in the design and development of AI assistants. Furthermore, individual differences, such as locus of control and future orientation, shape people's preferences toward AI technologies. Understanding these differences can inform the development of personalized AI systems that cater to individual needs. From a practical perspective, our findings can guide the development of AI systems that are not only efficient but also ethically sound and environmentally friendly.

Given the strong impact of cost on consumer preferences, relying only on consumer choice is inadequate for ensuring the ethical and environmentally sustainable development of AI systems, which highlights the importance of implementing robust governance strategies. Regulatory interventions and policy measures are essential to ensure AI development adheres to ethical standards and environmental sustainability independent of market forces.

**Table 4.** Conditional MMs on AI selections by future orientation

| | | | | 95% CI | |
|---|---|---|---|---|---|
| | MM | SE | p | Lower | Upper |
| **Present-oriented group:** | | | | | |
| **Satisfaction** | | | | | |
| 89% | 0.426 | 0.008 | <.000 | 0.410 | 0.442 |
| 94% | 0.510 | 0.007 | <.000 | 0.496 | 0.524 |
| 99% | 0.563 | 0.007 | <.000 | 0.549 | 0.578 |
| **Efficiency** | | | | | |
| 1h | 0.518 | 0.008 | <.000 | 0.503 | 0.534 |
| 3h | 0.505 | 0.007 | <.000 | 0.491 | 0.519 |
| 5h | 0.478 | 0.008 | <.000 | 0.463 | 0.493 |
| **Transparency** | | | | | |
| No transparency | 0.416 | 0.008 | <.000 | 0.400 | 0.431 |
| Low | 0.475 | 0.008 | <.000 | 0.460 | 0.490 |
| High | 0.611 | 0.008 | <.000 | 0.595 | 0.626 |
| **Costs** | | | | | |
| For free | 0.737 | 0.007 | <.000 | 0.723 | 0.751 |
| 300 yen | 0.488 | 0.008 | <.000 | 0.473 | 0.503 |
| 600 yen | 0.279 | 0.008 | <.000 | 0.264 | 0.294 |
| **Future-oriented group:** | | | | | |
| **Satisfaction** | | | | | |
| 89% | 0.423 | 0.009 | <.000 | 0.405 | 0.442 |
| 94% | 0.492 | 0.009 | <.000 | 0.475 | 0.509 |
| 99% | 0.583 | 0.009 | <.000 | 0.565 | 0.601 |
| **Efficiency** | | | | | |
| 1h | 0.547 | 0.010 | <.000 | 0.527 | 0.566 |
| 3h | 0.511 | 0.010 | <.000 | 0.492 | 0.530 |
| 5h | 0.446 | 0.010 | <.000 | 0.427 | 0.465 |
| **Transparency** | | | | | |
| No transparency | 0.388 | 0.009 | <.000 | 0.370 | 0.407 |
| Low | 0.452 | 0.009 | <.000 | 0.434 | 0.471 |
| High | 0.660 | 0.010 | <.000 | 0.641 | 0.679 |
| **Costs** | | | | | |
| For free | 0.690 | 0.010 | <.000 | 0.670 | 0.710 |
| 300 yen | 0.481 | 0.009 | <.000 | 0.463 | 0.498 |
| 600 yen | 0.323 | 0.010 | <.000 | 0.302 | 0.343 |

Proactive governance is crucial to direct AI development toward enhancing societal welfare, maintaining high ethical standards, and ensuring environmental sustainability as we advance into a more AI-driven society.

## Data availability

The materials and data in the present study are available at the OSF website: https://osf.io/hf7za/?view_only=0f5b30e1a7c5494a8e74e080232f39a2.

## Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

## Ethics approval

Ethical approval was obtained from Osaka University Ethics Committee.

## Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

## ORCID iDs

Tomohiro Ioku https://orcid.org/0000-0001-5499-6470
Jaehyun Song https://orcid.org/0000-0002-3692-6505

## Supplemental material

Supplemental material for this article is available online.

## References

Acikgoz Y, Davison KH, Compagnone M, et al. (2020) Justice perceptions of artificial intelligence in selection. *International Journal of Selection and Assessment* 28(4): 399–416.

Allen DG, Weeks KP and Moffitt KR (2005) Turnover intentions and voluntary turnover: The moderating roles of self-monitoring, locus of control, proactive personality, and risk aversion. *Journal of Applied Psychology* 90(5): 980–990.

Aoki N (2020) An experimental study of public trust in AI chatbots in the public sector. *Government Information Quarterly* 37(4): 101490.

Ballard K and Knutson B (2009) Dissociable neural representations of future reward magnitude and delay during temporal discounting. *Neuroimage* 45(1): 143–150.

Besley JC (2010) Public engagement and the impact of fairness perceptions on decision favorability and acceptance. *Science Communication* 32(2): 256–280.

Bijker W (1995) *Of Bicycles, Bakelites, and Bulbs: Toward a Theory of Sociotechnical Change*. Cambridge, MA: MIT Press.

Bruderer Enzler HB, Diekmann A and Liebe U (2019) Do environmental concerns and future orientation predict metered household electricity use? *Journal of Environmental Psychology* 62: 22–29.

Brügger A, Dessai S, Devine-Wright P, et al. (2015) Psychological responses to the proximity of climate change. *Nature Climate Change* 5(12): 1031–1037.

Buijsman S (2023) Navigating fairness measures and trade-offs. *AI and Ethics* 3.

Colquitt JA, Conlon DE, Wesson MJ, et al. (2001) Justice at the millennium: A meta-analytic review of 25 years of organizational justice research. *Journal of Applied Psychology* 86(3): 425–445.

Colquitt JA and Rodell JB (2015) Measuring justice and fairness. In: Cropanzano RS and Ambrose ML (eds) *The Oxford Handbook of Justice in the Workplace*. Oxford: Oxford University Press, 187–202.

Corral-Verdugo V and Pinheiro JQ (2006) Sustainability, future orientation and water conservation. *European Review of Applied Psychology* 56(3): 191–198.

Cowls J, Tsamados A, Taddeo M, et al. (2023) The AI gambit: Leveraging artificial intelligence to combat climate change—opportunities, challenges, and recommendations. *AI & Society* 38(1): 283–307.

Dauvergne P (2021) The globalization of artificial intelligence: Consequences for the politics of environmentalism. *Globalizations* 18(2): 285–299.

de Fine Licht K and de Fine Licht J (2020) Artificial intelligence, transparency, and public decision-making. *AI and Society* 35(4): 917–926.

Dietvorst BJ, Simmons JP and Massey C (2018) Overcoming algorithm aversion: People will use imperfect algorithms if they can (even slightly) modify them. *Management Science* 64(3): 1155–1170.

Dolata M and Schwabe G (2023) What is the metaverse, and who seeks to define it? Mapping the site of social construction. *Journal of Information Technology* 38(3): 239–266.

Earle TC and Siegrist M (2008) On the relation between trust and fairness in environmental risk management. *Risk Analysis* 28(5): 1395–1414.

Edwards L and Veale M (2017) Slave to the algorithm? Why a'right to an explanation' is probably not the remedy you are looking for. *Duke Law and Technology Review* 16: 18–84.

Engle-Friedman M, Tipaldo J, Piskorski N, et al. (2022) Enhancing environmental resource sustainability by imagining oneself in the future. *Journal of Environmental Psychology* 79: 101746.

*European Commission* (2021) *Declaration on a green and digital transformation of the EU*. Brussels: European Commission. Available at*:* https://digital-strategy.ec.europa.eu/en/news/eu-countries-commit-leading-green-digital-transformation (accessed 17 July 2023).

Fan X, Cai FC and Bodenhausen GV (2022) The boomerang effect of zero pricing: When and why a zero price is less effective than a low price for enhancing consumer demand. *Journal of the Academy of Marketing Science* 50(3): 521–537.

Felzmann H, Villaronga EF, Lutz C, et al. (2019) Transparency you can trust: Transparency requirements for artificial intelligence between legal norms and contextual concerns. *Big Data & Society* 6(1): 205395171986054.

Floridi L, Cowls J, Beltrametti M, et al. (2018) AI4People—An Ethical framework for a good AI society: Opportunities, risks, principles, and recommendations. *Minds and Machines* 28(4): 689–707.

Frederick S, Loewenstein G and O'Donoghue T (2002) Time discounting and time preference: A critical review. *Journal of Economic Literature* 40(2): 351–401.

Galvin BM, Randel AE, Collins BJ, et al. (2018) Changing the focus of locus (of control): A targeted review of the locus of control literature and agenda for future research. *Journal of Organizational Behavior* 39(7): 820–833.

Gold M (1999) *The Complete Social Scientist: A Kurt Lewin Reader*. Washington, DC: American Psychological Association.

Greenberg J (1990) Organizational justice: Yesterday, today, and tomorrow. *Journal of Management* 16(2): 399–432.

Grimmelikhuijsen S (2023) Explaining why the computer says no: Algorithmic transparency affects the perceived trustworthiness of automated decision-making. *Public Administration Review* 83(2): 241–262.

Gu D, Jiang J, Zhang Y, et al. (2020) Concern for the future and saving the earth: When does ecological resource scarcity promote pro-environmental behavior? *Journal of Environmental Psychology* 72: 101501.

Gunning D, Stefik M, Choi J, et al. (2019) XAI—Explainable artificial intelligence. *Science Robotics* 4(37).

Habuka H (2023) *Japan's Approach to AI Regulation and Its Impact on the 2023 G7 Presidency. (Report No. 14)*. Available at: https://www.csis.org/analysis/japans-approach-ai-regulation-and-its-impact-2023-g7-presidency.

Hainmueller J, Hopkins DJ and Yamamoto T (2014) Causal inference in conjoint analysis: Understanding multidimensional choices via stated preference experiments. *Political Analysis* 22(1): 1–30.

Hayashi Y (2007) The four perspectives on social justice: A review. *Japanese Journal of Social Psychology* 22(3): 305–330.

Henrich J, Heine SJ and Norenzayan A (2010) The weirdest people in the world? *Behavioral and Brain Sciences* 33(2–3): 61–83.

Hilton DJ (1990) Conversational processes and causal explanation. *Psychological Bulletin* 107(1): 65–81.

Hossain MT and Saini R (2015) Free indulgences: Enhanced zero-price effect for hedonic options. *International Journal of Research in Marketing* 32(4): 457–460.

Hüttel BA, Schumann JH, Mende M, et al. (2018) How consumers assess free E-services: The role of benefit-inflation and cost-deflation effects. *Journal of Service Research* 21(3): 267–283.

*IEA* (2020) *Data centres and data transmission networks*. Paris: International Energy Agency. Available at*:* https://www.iea.org/energy-system/buildings/data-centres-and-data-transmission-networks (accessed 19 July 2023).

Jobin A, Ienca M and Vayena E (2019) The global landscape of AI ethics guidelines. *Nature Machine Intelligence* 1(9): 389–399.

Joireman JA, Shaffer MJ, Balliet D, et al. (2012) Promotion orientation explains why future-oriented people exercise and eat healthy: Evidence from the two-factor consideration of future consequences-14 scale. *Personality and Social Psychology Bulletin* 38(10): 1272–1287.

Joireman JA, Van Lange PAM and Van Vugt M (2004) Who cares about the environmental impact of cars? Those with an eye toward the future. *Environment and Behavior* 36(2): 187–206.

Jones C, Hine DW and Marks ADG (2017) The future is now: Reducing psychological distance to increase public engagement with climate change. *Risk Analysis* 37(2): 331–341.

Kelley PG, Yang Y, Heldreth C, et al. (2021) Exciting, useful, worrying, futuristic: Public perception of artificial intelligence in 8 countries. *In: Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society*: 627–637.

Kieslich K, Keller B and Starke C (2022) Artificial intelligence ethics by design. Evaluating public perception on the importance of ethical design principles of artificial intelligence. *Big Data and Society* 9(1): 205395172210929.

Köbis N and Mossink LD (2021) Artificial intelligence versus Maya Angelou: Experimental evidence that people cannot differentiate AI-generated from human-written poetry. *Computers in Human Behavior* 114: 106553.

König PD, Felfeli J, Achtziger A, et al. (2022) The importance of effectiveness versus transparency and stakeholder involvement in citizens' perception of public sector algorithms. *Public Management Review* 26: 1–22.

König PD, Wurster S and Siewert MB (2022) Consumers are willing to pay a price for explainable, but not for green AI. Evidence from a choice-based conjoint analysis. *Big Data and Society* 9(1): 20539517211069630.

König PD, Wurster S and Siewert MB (2023) Sustainability challenges of artificial intelligence and Citizens' regulatory preferences. *Government Information Quarterly* 40(4): 101863.

Kordzadeh N and Ghasemaghaei M (2022) Algorithmic bias: Review, synthesis, and future research directions. *European Journal of Information Systems* 31(3): 388–409.

Krütli P, Stauffacher M, Pedolin D, et al. (2012) The process matters: Fairness in repository siting for nuclear waste. *Social Justice Research* 25(1): 79–101.

Landers RN and Behrend TS (2023) Auditing the AI auditors: A framework for evaluating fairness and bias in high stakes AI predictive models. *American Psychologist* 78(1): 36–49.

Lee CC, Qin S and Li Y (2022) Does industrial robot application promote green technology innovation in the manufacturing industry? *Technological Forecasting and Social Change* 183: 121893.

Lee MK, Jain A, Cha HJ, et al. (2019) Procedural justice in algorithmic fairness: Leveraging transparency and outcome control for fair algorithmic mediation. *Proceedings of the ACM on Human-Computer Interaction* 3(CSCW): 1–29.

Leeper TJ, Hobolt SB and Tilley J (2020) Measuring subgroup preferences in conjoint experiments. *Political Analysis* 28(2): 207–221.

Lenzi D, Balvanera P, Arias-Arévalo P, et al. (2023) Justice, sustainability, and the diverse values of nature: Why they matter for biodiversity conservation. *Current Opinion in Environmental Sustainability* 64: 101353.

Levenson H (1973) Multidimensional locus of control in psychiatric patients. *Journal of Consulting and Clinical Psychology* 41(3): 397–404.

Lünich M, Marcinkowski F and Kieslich K (2021) It's now or never! future discounting in the application of the online privacy calculus. *Cyberpsychology: Journal of Psychosocial Research on Cyberspace* 15(3): Article 11.

Malgieri G and Comandé G (2017) Why a right to legibility of automated decision-making exists in the general data protection regulation. *International Data Privacy Law* 7(4): 243–265.

Marcinkowski F, Kieslich K, Starke C, et al. (2020) Implications of AI (un-)fairness in higher education admissions: The effects of perceived AI (un-)fairness on exit, voice and organizational reputation. *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency*: 122–130.

McClure PK (2018) 'You're fired,' says the robot: The rise of automation in the workplace, technophobes, and fears of unemployment. *Social Science Computer Review* 36(2): 139–156.

McFarlin DB and Sweeney PD (1992) Distributive and procedural justice as predictors of satisfaction with personal and organizational outcomes. *Academy of Management Journal* 35(3): 626–637.

Miller T (2019) Explanation in artificial intelligence: Insights from the social sciences. *Artificial Intelligence* 267: 1–38.

*Ministry of internal affairs and communications* (2022) *Further Promotion of Safe, Secure, and Trustworthy Implementation of AI in Society*. Available at: https://www.soumu.go.jp/iicp/research/results/ai-network.html (accessed 19 July 2023).

Mukai T, Yuyama Y and Watamura E (2023) Support for using AI in trials and its determinants. *Joho Chishiki Gakkaishi* 33(1): 3–19.

Ng TWH, Sorensen KL and Eby LT (2006) Locus of control at work: A meta-analysis. *Journal of Organizational Behavior* 27(8): 1057–1087.

Niemand T, Mai R and Kraus S (2019) The zero-price effect in freemium business models: The moderating effects of free mentality and price–quality inference. *Psychology and Marketing* 36(8): 773–790.

Nussberger A-M, Luo L, Celis LE, et al. (2022) Public attitudes value interpretability but prioritize accuracy in artificial intelligence. *Nature Communications* 13(1): 5821.

Ohnuma S, Yokoyama M and Mizutori S (2022) Procedural fairness and expected outcome evaluations in the public acceptance of sustainability policy-making: A case study of multiple stepwise participatory programs to develop an environmental master plan for Sapporo, Japan. *Sustainability* 14(6): 3403.

Palmeira MM and Srivastava J (2013) Free offer ≠ cheap product: A selective accessibility account on the valuation of free offers. *Journal of Consumer Research* 40(4): 644–656.

Pinch TJ and Bijker WE (1984) The social construction of facts and artefacts: Or how the sociology of science and the sociology of technology might benefit each other. *Social Studies of Science* 14(3): 399–441.

Prahl A and Van Swol L (2017) Understanding algorithm aversion: When is advice from automation discounted? *Journal of Forecasting* 36(6): 691–702.

Promberger M and Baron J (2006) Do patients trust computers? *Journal of Behavioral Decision Making* 19(5): 455–468.

Rotter JB (1954) *Social Learning and Clinical Psychology*. Englewood Cliffs, NJ: Prentice Hall, Inc.

Rotter JB (1966) Generalized expectancies for internal versus external control of reinforcement. *Psychological Monographs* 80(1): 1–28.

Ryff CD (1989) Happiness is everything, or is it? Explorations on the meaning of psychological well-being. *Journal of Personality and Social Psychology* 57(6): 1069–1081.

Samek W and Müller K-R (2019) *Towards Explainable Artificial Intelligence BT—Explainable AI: Interpreting, Explaining and Visualizing Deep Learning* Samek W, Montavon G, Vedaldi A et al. (eds). Springer International Publishing, pp. 5–22. https://doi.org/10.1007/978-3-030-28954-6_1.

Schiff DS, Schiff KJ and Pierson P (2022) Assessing public value failure in government adoption of artificial intelligence. *Public Administration* 100(3): 653–673.

Schindler R (2011) *Pricing strategies: A marketing approach*. Sage Publications, Inc.

Shampanier K, Mazar N and Ariely D (2007) Zero as a special price: The true value of free products. *Marketing Science* 26(6): 742–757.

Shin D (2020) User perceptions of algorithmic decisions in the personalized AI system: Perceptual evaluation of fairness, accountability, transparency, and explainability. *Journal of Broadcasting and Electronic Media* 64(4): 541–565.

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: 102551.

Shin D, Lim JS, Ahmad N, et al. (2022) Understanding user sensemaking in fairness and transparency in algorithms: Algorithmic sensemaking in over-the-top platform. *AI and Society* 39(2): 477–490.

Shin D and Park YJ (2019) Role of fairness, accountability, and transparency in algorithmic affordance. *Computers in Human Behavior* 98: 277–284.

Shove E and Pantzar M (2005) Consumers, producers, and practices: Understanding the invention and reinvention of Nordic walking. *Journal of Consumer Culture* 5(1): 43–64.

Shove E and Walker G (2010) Governing transitions in the sustainability of everyday life. *Research Policy* 39(4): 471–476.

Song J (2022) q2c: Qualtrics to conjoint. R package. Available at: https://github.com/JaehyunSong/q2c#readme.

Spector PE (1982) Behavior in organizations as a function of employee's locus of control. *Psychological Bulletin* 91(3): 482–497.

Spiegel U, Benzion U and Shavit T (2011) Free product as a complement or substitute for a purchased product–does it matter. *Modern Economy* 02(2): 124–131.

Starke C, Baleis J, Keller B, et al. (2022) Fairness perceptions of algorithmic decision-making: A systematic review of the empirical literature. *Big Data and Society* 9(2): 20539517221115188.

Stead WW (2018) Clinical implications and challenges of artificial intelligence and deep learning. *JAMA* 320(11): 1107–1108.

Strathman A, Gleicher F, Boninger DS, et al. (1994) The consideration of future consequences: Weighing immediate and distant outcomes of behavior. *Journal of Personality and Social Psychology* 66(4): 742–752.

Strubell E, Ganesh A and McCallum A (2019) Energy and policy considerations for deep learning in NLP. *ArXiv Preprint ArXiv:1906.02243*.

Sunshine J and Tyler TR (2003) The role of procedural justice and legitimacy in shaping public support for policing. *Law and Society Review* 37(3): 513–547.

Tham YJ, Hashimoto T and Karasawa K (2022) Underlying dimensions of benefit and risk perception and their effects on people's acceptance of conditionally/fully automated vehicles. *Transportation* 49(6): 1715–1736.

Thibaut J and Walker L (1975) Procedural justice: A psychological analysis. Erlbaum. Available at: https://cir.nii.ac.jp/crid/1130000797754273408.bib?lang=ja.

Trope Y and Liberman N (2010) Construal-level theory of psychological distance. *Psychological Review* 117(2): 440–463.

Tyler TR (2000) Social justice: Outcome and procedure. *International Journal of Psychology* 35(2): 117–125.

Tyler TR and Caine A (1981) The influence of outcomes and procedures on satisfaction with formal leaders. *Journal of Personality and Social Psychology* 41(4): 642–655.

Tyler TR, Goff PA and MacCoun RJ (2015) The impact of psychological science on policing in the United States: Procedural justice, legitimacy, and effective law enforcement. *Psychological Science in the Public Interest* 16(3): 75–109.

Vinuesa R, Azizpour H, Leite I, et al. (2020) The role of artificial intelligence in achieving the sustainable development goals. *Nature Communications* 11(1): 33.

Wakslak C and Trope Y (2009) The effect of construal level on subjective probability estimates. *Psychological Science* 20(1): 52–58.

Wang Q, Bowling NA and Eschleman KJ (2010) A meta-analytic examination of work and general locus of control. *Journal of Applied Psychology* 95(4): 761–768.

Watamura E, Ioku T, Mukai T, et al. (2023) Empathetic robot judge, we trust you. *International Journal of Human-Computer Interaction* 40: 1–10.

Yeung K (2018) Five fears about mass predictive personalization in an age of surveillance capitalism. *International Data Privacy Law* 8(3): 258–269.

Yokoi R and Nakayachi K (2024) Human motivation in competition against artificial intelligence: Using one-to-one games. *International Journal of Human–Computer Interaction* 40: 1–13.

Zhang X, Grisolía JM and Lane TDT (2023) Zero price effect on hotel demand: Evidence from a discrete choice experiment. *Tourism Management* 96: 104692.

Zimmerman MA and Rappaport J (1988) Citizen participation, perceived control, and psychological empowerment. *American Journal of Community Psychology* 16(5): 725–750.