Erasmus University Rotterdam

Erasmus School of Economics

Bachelor Thesis Marketing

Can (A)I Change Your Mind?

Author: Pool, Devin

Student Number: 561017

Supervisor: Van Hasselt, M.J.L.

Second Reader: Dr. D. Zilbershtein

Finish Date: Aug. 15th, 2024

Acknowledgments

I would like to express my gratitude to my professor and thesis supervisor, Mr. Van Hasselt, for his feedback, guidance, and especially patience through the thesis process. His support and expertise were essential in helping navigate through the complexities of consumer behaviour and artificial intelligence.

Additionally, I would like to thank my friends and family for their moral support, assistance in the early stages of ideation, and voluntary participation in interviews that provided invaluable information that proved to be essential building blocks that the thesis was built upon. A final thanks goes to those who participated in the online experiment.

Table of Contents

Abstract

During today's exponential growth of the AI industry, many questions are asked regarding their potential to replace humans in professional roles. This paper investigates the differences in effectiveness between humans and artificial intelligence as recommending agents. By exploring this relationship through a randomized online experiment ($N = 270$), the author reveals the agents' capabilities in the niche scenario known as a preference reversal. The data was analyzed using statistical tests, including a one-sample t-test to estimate whether a preference reversal was possible and a two-way ANOVA to explore the moderating effect of product type. The main findings shows that both recommending agents were capable of reversing preferences, while human recommenders had a significant advantage for scenarios involving a hedonic product. Despite this, the field of consumer behavior remains a vastly complex one, and the conclusions from this paper would largely benefit from future research that investigates other variables that may also be responsible for reversing preferences.

1 Introduction

The recent boom in the popularity of artificial intelligence, largely attributed to the success of ChatGPT, has flooded the world with questions about what is achievable with AI. While some were concerned that AI would lead to lost jobs, others sought ways to benefit from this innovation. The author aims to fit into the latter category by asking how AI could play a role in the modern-day decision-making process, specifically in helping people make choices that truly satisfy their needs.

Past and present research suggests differences in people's perceptions of advice given by humans and machines. Human recommenders commonly take the form of other consumers or sales agents. While other consumers are often perceived as honest sources of information with no ulterior motive behind their advice, sales agents could be seen as people who merely want to maximize their revenue by selling you the more expensive option (Wien & Peluso, 2021). AI recommenders are a relatively new concept compared to their human counterparts and create recommendations based on past consumer data. As the process of recommendation is automated through algorithms, people do not perceive AI as agents who are capable of selfish intent, unlike humans (Kim & Duhachek, 2020). Despite this, people still tend to avoid advice from machines, a phenomenon known as algorithmic aversion (Dietvorst et al., 2015). The AI field is now rapidly developing and becoming more accessible to regular people than ever before, and as its popularity increases, people's familiarity with the technology helps to dwindle algorithmic aversion (Logg et al., 2019). Thus, as trust in artificial intelligence grows, the question arises: would AI be more effective as a recommender than its human counterpart?

To answer this question, the author explores the decision-making process and how preferences can change. The decision-making process can be simplified into four stages: problem definition, information seeking, acting on the decision, and post-purchase evaluation (Hoyer et al., 2017). Internal and external influences, such as one's perception of the world and surrounding culture, continue to shape an individual's preferences beyond the initial stage of the decision-making process (Hoyer et al., 2017). Despite several studies on preference formation, people are commonly unaware of what causes them to make certain decisions and regularly choose options that do not maximize their utility (Ariely, 2010; Donkers, 2013). The author specifically chose to explore the concept of preference reversals to contribute new findings to the current understanding of preference manipulation. This specifically investigates the role of recommenders in changing a person's preferences (e.g. from initially preferring product *A* over *B,* to now preferring product *B* over *A*).

Evidently, consumer behavior is complex. The author does not aim to cover all the topics and nuances of what variables contribute to certain behaviors and decisions, as it is beyond the scope of this study. Instead, the author aims to provide evidence of the effectiveness of recommending agents in their ability to convince consumers to choose products they *should* prefer. Furthermore, the author discusses the moderating role of product type (hedonic versus utilitarian) to gain insight into how recommender types can have an advantage over others when considering different products. This may be especially insightful for practical applications of recommender types, where businesses would be able to explore which recommender is best suited for the products they want to sell.

2 Theoretical Framework

2.1 Preferences

2.1.1 Defining Preference

Consumer behavior studies choices consumers make when searching, evaluating, purchasing, and using products/services they believe would satisfy their wants and needs (Schiffman & Wisenblit, 2018). Within the vast study of consumer behavior lies the subject of consumer preferences, defined as *"the fact of people liking or wanting one thing more than another"* (Cambridge Business English Dictionary, 2024). Slovic (1995), examined how preference is constructed and described its expression as "the essence of intelligent, purposeful behavior".

Adam Smith first introduced the concept of rational choice theory in 1776, which closely relates to preference. The general understanding of this theory suggests that an individual's decisions are deliberate, consistent, and driven by some rationale, such as utility maximization (Edwards, 1954; Ulen, 1999). There exist two fundamental assumptions about preferences: completeness and transitivity. Completeness suggests that all alternatives are comparable, such that an individual will never exclude an alternative in their decision-making process. Transitivity suggests that preference is transitive; for a set of alternatives *x*, y, and *z*, if *x* is preferred to *y* and *y* is preferred over *z*, this shows that *x* has to be preferred over *z* (Ulen, 1999). Moreover, preferences are not confined to a single degree; instead, they can be categorized into three categories: a weak preference, a strict preference, or an indifference (Strohmaier & Messerli, 2024).

Adding to the complexity of preferences is the research indicating potential differences in what individuals *say* they prefer versus their actual preferences adds to the complexity of preferences (Wardman, 1988). This difference in the "stated preference" and "real preference" can be attributed to a hypothetical bias; individuals reporting unrealistic choices under hypothetical/test conditions (Ajzen et al., 2004; De Corte et al., 2021). Despite the possibility that the stated preferences do not fully capture real preferences, they are a common proxy used in marketing research and more appropriate for this thesis due to an ease of measurement (Wardman, 1988).

2.1.2 Decision-Making Process

The decision-making process can be summarized into four stages: problem recognition, information search, acting on the decision, and post-purchase evaluation (Hoyer et al., 2017). Though seemingly straightforward, the factors that influence each stage of the process create a vastly complex network, one that cannot be fully explored within the scope of the paper.

To briefly touch on the topic, the influencing factors are commonly grouped into two categories: internal and external influences. Internal influences relate to the consumer's psychological factors such as self-concept (ability, motivation, and opportunity), perception, comprehension, and knowledge. External influences are the factors that make up the consumer's surroundings, such as culture, social class, family roles, and reference groups (Hoyer et al., 2017). The significance of each factor differs per individual, but understanding how these factors could affect an individual is an essential component of consumer behavior studies.

As many factors are responsible for influencing consumer decisions, this thesis will address the issue of controlling for these variables through a randomized controlled trial, where individuals are randomly assigned one of four test conditions. The random assignment will ensure an equal distribution of these variables, eliminating systematic differences that may skew the data as much as possible within the constraints of this thesis.

As shown, decision-making and the formation of preferences are complicated processes. It is so complex that many newer behavioral researchers argue that the traditional view of the process is systematically biased and flawed, as many real human decision-making are "predictably irrational" (Ariely, 2010; Gilovich et al., 2002). This opens up the conversation about how people may require assistance to optimally reach their preferred state.

2.2 Changing Preferences

2.2.1 Defining Preference Change

A preference change refers to the alteration of a person's choices or desires, driven by new information, motivation, or context (Donkers, 2013; Strohmaier & Messerli, 2024).

A preference change can take one of four basic forms: revision, contraction, addition, or subtraction (Hansson, 1995). While more complex forms of preference change exist, they are built on these foundational forms. This paper will examine a preference reversal, a sub-type of preference revision: a change in preference within a set of alternatives without the addition or subtraction of new alternatives (Hansson, 1995). Specifically, a preference reversal occurs when an individual initially prefers one option over another but reverses that preference due to a change in the decision-making environment (Johnson et al., 1988). An example of a preference reversal is when a voter initially prefers political Party A over Party B. If Party A were to change its stance on a key issue, the voter may revise their preference and now prefer Party B. An example of this phenomenon is demonstrated in Figure 1. Note that changes in the decision-making environment that influences the "new preference" can be a variety of things, what will be discussed in this paper specifically is a recommending agent (AI and human).

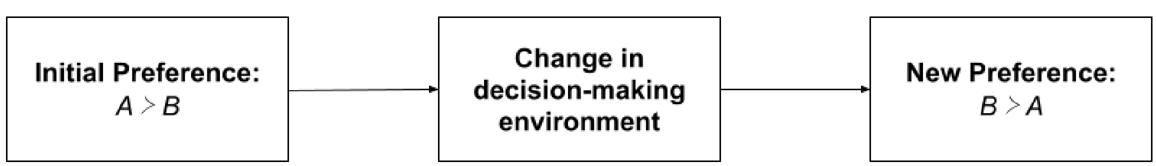


Figure 1. Preference Reversal Demonstration

The concept of “preference change” has sparked debates as it suggests preferences are not consistent as this notion challenges the core assumptions of the rational choice theory (Johnson et al., 1988). However, McKenzie (2018) suggest that preference reversals do *not* violate any of the original rational choice assumptions under specific conditions. The researchers clarify a phenomenon they refer to as a joint-separate preference reversal. This occurs when preference changes due to a change in the evaluation context, either jointly or separately evaluated. For example, in an evaluation of restaurants *A* and *B*, when rating them separately, an individual may rank restaurant *A* higher than restaurant *B*. However, when presenting both options together, allowing for direct comparisons, the preference may shift towards restaurant B. This scenario provides grounds for believing that preferences are dynamic and capable of updating based on changes in the decision-making environment.

2.2.2 Key Studies

Donkers (2013), showed that consumers are often not rational in their decision-making, often neglecting, forgetting, or simply unaware of important factors that objectively determine which alternative would be the optimal choice – contributing the greatest amount of utility. He attributed this to consumers being heavily influenced by the context they are in when making their decision, and if the context were to change, preferences would change along with it. Donkers presented that changes in the salience of consequences from a decision would ultimately affect a person's preferences. This was demonstrated with an example of a person’s retirement-saving habits; once the individual was shown the consequences of his current high-spending habits—his future self becoming less happy—causing the individual to alter his habits to improve his future outcome. This study showcases that preferences are able to change through manipulating the salience of consequences.

Sher and McKenzie (2014), discuss the role of context in changing preferences. Traditional principles of rational choice indicate that preferences should be coherent (logical and consistent) and suggesting that they are always static. This principle implicitly assumes that context does not provide relevant information for decision-making, thus claiming that preferences cannot change in different contexts (Sher & McKenzie, 2014). The researchers ultimately challenged this notion, showing that preferences are in fact dynamic and claiming that the dynamic nature of preferences is an integral part of the human ability to adapt to different situations. Similar to the conclusion drawn by Donkers (2013), the researchers revealed how context serves as a cue in to help people during the decision-making, particularly in scenarios with limited information and prior knowledge.

Another form of preference change through context is how humans develop and change over time. Whitson et al. (2014) conducted a follow-up experiment to one they completed in 1996 to explore the utility consumers gain from certain product attributes. After examining the data, the study revealed that the utility obtained from the attributes had significant differences from the results of the 1996 experiment. A notable finding from this experiment is that in 2014, consumers deemed it more acceptable to pay higher prices for eco-friendly products compared to 1996. The researchers believe that consumers' conditioning to believe eco-friendly processes are more expensive is responsible for this change in preference towards prices. This experiment demonstrates a shift in preferences and beliefs over the course of 18 years.

As evident from the discussed experiments and contrary to traditional economic assumptions (that humans are rational agents), real-world behavior often deviates from theorized rationality, as nonoptimal decisions are common (Ariely, 2010). This discrepancy highlights the potential role of a recommender—whether a person or a machine—who can provide guidance based on consumers' real preferences. Such recommendations can help consumers make decisions that more closely align with their “true preferred state” by highlighting aspects that may be initially overlooked or undervalued.

2.3 Recommender Types (AI vs. Human)

2.3.1 AI Recommenders

AI recommendation systems are algorithms that make use of statistical and actuarial models to analyze consumer data and predict preferences. By leveraging large datasets, these algorithmic models can create personalized recommendations for relevant consumers (Grove & Meehl, 1996). AI recommenders can take several forms and do not require direct communication with the consumer; instead, they could take the form of a list of products that the consumer may find appealing (Häubl & Murray, 2003). AI recommenders that directly interact with individuals can have a virtual body, face, and voice that either imitate human-like qualities or intentionally appear more robotic; these are commonly known as AI chatbots (Araujo, 2018; Mende et al., 2019).

Today, companies use algorithmic recommendation systems to suggest products/services for consumers based on their available data (Budzinski et al., 2021; Kim et al., 2021). By implementing these AI recommendation systems, companies have a self-improving system that analyzes consumer data, outputs personalized recommendations, observes how those recommendations perform, and re-evaluates itself based on the performance (Budzinski et al., 2021). The recent worldwide mass adoption of AI technology for businesses (and independent use) proves that the technology benefits society, thus enforcing the importance of understanding how this innovation can be used optimally (Kim et al., 2021).

This paper conceptualizes the AI recommender as an AI chatbot without a virtual body, face, or voice. The focus of the research is on how the consumer interprets the advice given; thus, the consumers will have no opportunity to communicate with the AI recommender other than receiving the recommendation.

2.3.2 Human Recommenders

Human recommenders can fall into two major categories: another consumer (friends, family, relatives, or strangers) or an expert (sales-representative or independent experts) (Senecal & Nantel, 2004). Human recommenders fitting into the “another consumer” category, advise on products/services as a form of altruistic behavior and are commonly driven by empathic concern for others (Hennig-Thurau et al., 2004). However, the “expert” category provides advice for self-serving purposes such as economic incentives or the desire to earn respect/recognition as an expert in that field (Constant et al., 1996; Sah & Loewenstein, 2014. Research demonstrates that recommenders, regardless of their type, play a crucial role in marketing and can significantly influence a company's sales (De Bruyn & Lilien, 2008). Çelen et al. (2010) conducted a study showing that individuals acted on a recommendation 74% of the time.

In this research paper, the author defines the human recommender as a previous customer offering digital advice. The human recommender, like the AI recommender, will not engage in active communication with the consumer; instead, they will display a one-off advice as digital on-scereen text. The human recommender will not have any ulterior motives for recommending the product, leading to a scenario where the consumer receiving the advice tends to trust it (Boerman et al., 2017).

2.3.3 AI vs. Human

AI recommenders share similarities with commercial agents, as they operate under the constraints and instructions of the company deploying them. Yet consumers typically do not view AI recommenders as commercial agents attempting to persuade them to make a particular purchase (Kim & Duhachek, 2020). Instead, consumers view AI as a tool that can help them make more informed decisions (Senecal & Nantel, 2004). Consumers are seen to be more comfortable disclosing private information to AI agents compared to their human counterparts, indicating differences in the perception of trustworthiness between AI and humans (Kim et al., 2021). The overall differences in how consumers interact with the two recommender types suggest that they are perceived as separate sources of information rather than alternatives (Wien & Peluso, 2021).

In terms of predicting events and outcomes, algorithmic systems can objectively outperform human judgments and intuitions due to their greater data processing capabilities (Dawes, 1979; Sanders & Manrodt, 2003). Despite its proven superiority in forming calculated judgments, people commonly resist algorithmic advice. Furthermore, research indicates that people weigh human input higher than algorithmic input (Önkal et al., 2009) and more harshly judge professionals who make use of algorithmic advice (Shaffer et al., 2012). This phenomenon is known as algorithm aversion (Dietvorst et al., 2015).

Contrastingly, Logg et al. (2019) argue that the prior beliefs of algorithmic aversion may be outdated as people increasingly use algorithms to help make informed decisions. The researchers proved that “algorithmic appreciation” exists for certain contexts. They showed that in six experiments, for tasks involving the estimation and prediction of events, individuals significantly preferred the advice of algorithms to that of humans and even to their personal judgment (although less prevalent). Paradoxically, when conducting a similar experiment with people who were considered experts in their field, these individuals relied less on advice from the algorithms compared to the non-experts.

2.3.4 Context

How people use advice received from algorithms or humans is shown to be dependent on the context in which the advice is given. Jin and Zhang (2023) discovered that people perceive AI recommendations as more competent in situations involving material purchases, while preferring human recommendations for experiential purchases like vacations or movie viewings. It has been shown that individuals tend to avoid AI for subjective and intuition-involved choices while accepting AI advice for analytical tasks (Kim et al., 2021). This will be extensively discussed in the upcoming section on hedonic and utilitarian products (Section 2.5).

Wang et al. (2022), explored the circumstances that affected individuals' preference for AI recommendations over their judgment. The researchers concluded that when the "stakes" are higher, humans' trust in AI recommendations dwindles, opting to trust their personal judgment instead. This shows it is not merely the nature of the good that affects human's trust in AI, but also the perceived cost of their decision. To reduce the prevalence of this effect in the experiment, this paper will present all decision scenarios with both recommender types and attempt to have all scenarios be of equal perceived cost.

2.3.5 Language Type

Similar to the role of AI recommenders in steering consumer preferences are 'Virtual Influencers', computer-generated digital characters taking on the role and personality of a social media influencer guided by a team (Ozdemir et al., 2023). Notably, both virtual influencers and AI recommenders face the same kind of mistrust: dehumanization, the belief that machines are incapable of emotion and thus not suitable as a source for matters requiring emotional thinking (Castelo et al., 2019). However, Ozdemir et al. (2023) show that the dehumanization of virtual influencers subsides depending on the language they use to communicate their message. When they use rational endorsement language, as opposed to emotional language, people perceive them as just as credible and effective as their human counterparts at promoting a brand or product.

Given that language type influences how humans perceive recommenders, the experiment will control for this by giving the same advice and recommendations for both recommenders. The visible difference in the presentation of information variable will be the source of the information, where one scenario would indicate the source as "customer" and the other as "AI".

2.3.6 Summary of Recommender Types

For this paper, AI recommenders are conceptualized as digital chatbots. These agents create recommendations by analyzing past consumer data. Human recommenders are defined as past consumers who present advice based on their personal experiences. As discussed, the variable of language type is known to be influence people's perception of virtual influencers and may be used to decrease the "dehumanization" that is related to algorithmic aversion. To avoid any effects from this variable, the author opted to have both recommender types exhibit the same recommendation message (Appendix A). With the distinctions between the recommender types made clear, the following section will present past studies that examined the relationship between preferences and recommender types.

2.4 Preferences and Recommenders

Allen (1953) investigated how the credibility of recommenders influences people's opinions. The recommenders were classified as either 'trustworthy' or 'untrustworthy.' The study measured how much information participants acquired, the effectiveness of the recommenders in changing opinions, and the retention of information. The results showed no significant differences between the two types of recommenders in terms of participant information acquisition and retention. However, the 'trustworthy' recommender had a significantly higher success rate in changing opinions compared to the 'untrustworthy' recommender. This demonstrates that recommender credibility plays a crucial role in altering people's opinions, which is similar to the preference reversal scenario this paper aims to explore. To eliminate the effect of differing credibility, this paper's experiment will present both AI and human recommenders as equally credible sources.

A study by Longoni and Cian (2020), explored how individuals perceived the competency of AI and humans as recommenders. Furthermore, they explored the moderating effects of hedonic and utilitarian attributes of products on perceived competency. The researchers hypothesized the "word-of-machine" effect; whereby people interpret advice from AI as being more competent for objects of a utilitarian nature. In their experiment, participants were primed with either a hedonic or utilitarian goal before interacting with recommenders. Once a goal was primed, they were given recommendations from either a human or AI recommender, followed by a decision about their preferred product based on the information provided. The researchers concluded that consumers preferred AI recommendations for utilitarian goals and human recommendations for hedonic goals. Furthermore, the researchers concluded that advice by AI recommenders was able to effectively highlight the utilitarian qualities of a product that would otherwise be overlooked, with a similar conclusion for the relationship between human recommenders and the hedonic qualities of a product.

This paper aims to build upon these previous findings by exploring how AI, or human recommenders, can influence people to alter their preferences by exposing them to a recommending agent. The author aims to observe a niche scenario, known as a preference reversal, which is yet to be explored in the context of AI versus human recommenders. Findings from this study will allow for a deeper understanding of how consumers interact with and interpret recommendations from different types of recommenders. Donkers (2013) demonstrates that the context of a consumer's decision heavily influences their choice; therefore, the author hypothesizes that exposing consumers to new information through recommenders creates a new context, enabling a re-evaluation of their preferences and potentially leading to a preference reversal.

H1*: Consumers will reverse their preferences after being exposed to a recommendation by either recommender type*

2.5 Hedonic vs. Utilitarian Products

As briefly discussed in the Longoni and Cian (2020) experiment, the hedonic and utilitarian qualities of a product can affect the perceived competency of recommender types. With those findings in mind, this paper aims to examine how product types will affect the effectiveness of recommender types in inducing a preference reversal.

2.5.1 Defining Hedonic and Utilitarian Products

The hedonic products are defined as giving sensory or experiential pleasure, such as going to the cinema, theme park, or vacation. The value of hedonic products is often derived based on experiential, emotional, and sensory dimensions (Botti & McGill, 2011). Conversely, utilitarian products serve a specific purpose, like monetary gain, or offer functional utility. The value of this product type is commonly evaluated on factual, rational, and logical dimensions (Longoni & Cian, 2020 Emotions often drive hedonic consumption, whereas cognitive factors drive utilitarian products (Batra & Ahtola, 1991; Botti & McGill, 2011. Despite the fact that products frequently combine both hedonic and utilitarian elements, this paper will categorize product types based on their primary characteristic: hence a hedonic or a utilitarian product.

2.5.2 Relation to Recommenders and Preferences

People's preconceived beliefs of AI systems as being unempathetic, factual, and analytical are similar to the value drivers of utilitarian products, resulting in individuals perceiving AI as a more competent source for utilitarian recommendations. Contrastingly, hedonic products are primarily emotionally driven, which has similarities with the "human" aspect of being empathetic, emotional, and capable of affective experiences. This preconception creates an association between human recommendations and hedonic values, making humans seemingly "better suited" to advise on hedonic products. This indicates that the optimal choice of recommender type plays a role in influencing a consumer's preference, depending on the type of product in the context.

This paper explores the moderating role of hedonic and utilitarian products in the context of recommender types and preference reversals. Based on the findings of Longoni and Cian (2020), the author hypothesizes a similar effect on preference reversals, where the advice from human recommenders is more effective in creating preference reversals for hedonic products than that of AI for utilitarian products.

H2*: Human recommenders cause significantly higher EPR than AI recommenders when considering hedonic products*

H3*: AI recommenders cause significantly higher EPR than human recommenders when considering utilitarian products*

2.6 Conceptual Framework

The figure below summarizes the discussed topics and formulated hypotheses. The author hypothesized that the product type would moderate the effect of recommender type on preference reversals. Specifically, both recommender types will be successful in inducing a preference reversal (H1). While the product type will moderate this effect, a hedonic (utilitarian) product evaluation will be more susceptible to a preference reversal when recommended by a(n) human (AI) (H1 & H2).

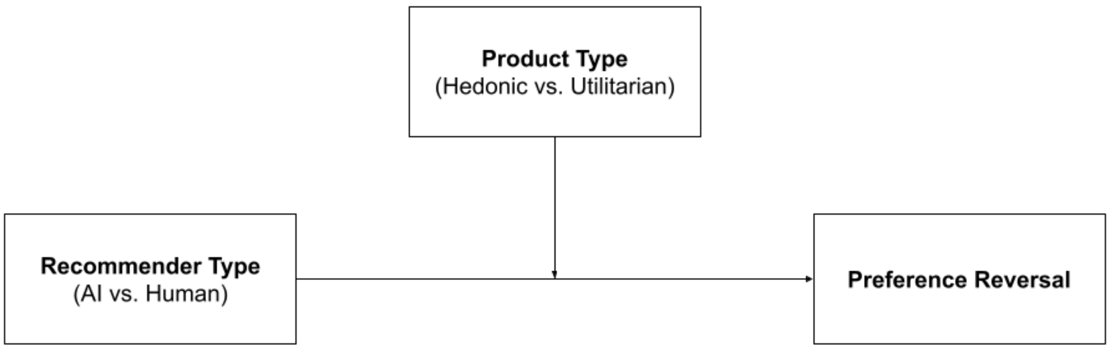


Figure 2. Conceptual Framework

3 Methodology

3.1 Research Method

To investigate the effectiveness of the recommender types in causing a preference reversal, the author conducted a quantitative study in the form of a randomized online experiment. This thesis also aims to investigate the potential moderating effects of product type, hedonic versus utilitarian. A quantitative approach is required to observe changes in the relationship between the predictor and outcome variables. The complex relationship between the recommenders, product types, and preference reversals can be better analyzed and represented using statistical tests.

3.2 Experiment Design

The experiment was designed entirely on Qualtrics, including the automated random assignment of test conditions. The experiment featured four randomized conditions, each consisting of two hedonic or utilitarian products (A and B), and a recommendation from either human or AI agents to influence their preference (Appendix A). This 2 x 2 experiment design was similar to the experiments conducted by Longoni and Cian (2020). A between-subject design was chosen as it allows for a faster completion time; eliminating potential carryover and fatigue effects that may be caused by a prolonged experiment. The products representing the hedonic and utilitarian categories were a pair of headphones and a laptop, respectively, chosen as these were successfully used in a similar experiment (Wien & Peluso, 2021). Each product presentation included a comparison table of product attributes that were used to imitate the information-seeking stage of the decision-making process. The test condition was further manipulated by the framing of a utilitarian/hedonic goal to guarantee the participant's understanding of the purchasing context. Participants were asked to state their "initial preference" *(INP)* on a 7-point Likert scale from "strongly prefer product A" to "strongly prefer product B". They were then randomly assigned a recommender (AI or human) to recommend the *opposing* product in an attempt to reverse their preference. The participants were to state if their preferences had changed, if so, they were shown a new Likert scale to specify their "new preferences" *(NWP)*. The difference between the *INP* and *NWP* variables was used to create the "Effective Preference Reversal" *(EPR)* variable to measure the effectiveness of each recommender in causing a preference reversal. As discussed in the theoretical framework, language type, recommender credibility, and perceived cost are extraneous variables that may influence the participants' preferences. To control for their potential effect on the outcome variable, these variables were kept constant for all test conditions.

3.3 Sampling

The goal for the sample was to collect a minimum of 250 responses. The online experiment remained active between July 10 - August 4th, 2024. The experiment was created in English and distributed on various social media platforms accessible by the author (Instagram, WhatsApp, Discord, and X). Additionally, participants were asked to share the experiment with others they knew. The sample collection used was a mixture of convenience and snowball sampling, resulting in a non-probability method. The author acknowledges potential validity issues with the chosen sampling methods, especially regarding homogeneity, selection bias, and limited generalizability. Given the available resources, the author chose to use these sampling methods despite their disadvantages. The author incorporated stages of randomization within the experiment to alleviate potential issues stemming from the sampling method.

The population studied for this thesis are young adults in Europe within the age group of 18-29, appropriately chosen due to their familiarity with online purchases, digital recommenders, and AI systems (Kennedy et al., 2023). The chosen age group's extensive exposure to digital technology makes them ideal candidates for testing the potential effects of the emerging AI innovation as recommenders.

3.4 Data Analysis

The collected data was analyzed using a two-way ANOVA (analysis of variance), one-sample t-test, a two-sample t-test, and the Wilcoxon-signed rank test, which were all conducted on Stata/MP 18.

The one-sample t-test is used to examine if the recommenders were able to induce preference reversals by testing if the mean *EPR* (effective preference reversal) significantly differs from zero. For added confidence of the conclusion of this one-sample t-test, a Wilcoxon-signed rank test will be used to test if the median EPR also differs from zero, chosen because it is robust to non-normal data. The findings of these tests will give an answer to H1.

The two-way ANOVA is used to partially answer H2 and H3 by showing the main and interaction effects between the recommender type and product type on *EPR*. To fully answer the hypotheses, a two-sample t-test was used to calculate if each recommender type was more effective than the other for the specific product types.

Before conducting any hypothesis testing, the author ensured that the data was representative of the target population and adhered to the assumptions of the statistical tests. The tested assumptions were normality and homogeneity of variances. Additionally, the author tested if the random assignment in the experiment was successful.

4 Results

4.1 Descriptive Statistics

The online experiment resulted in 270 observations. The respondents were 50.97% male, 47.58% female, and 1.49% identified as other. As the method used was non-probability sampling through social media channels accessible by the author (who is based on West Europe), this ensures that the majority of participants are from the Western Europe region. The sum of the three most common age groups (18-21, 22-25, 25-29) made up 94.07% of the total observations, hence representative of the target age group of 18-29. The highest obtained education observed was a Bachelor’s degree (47.58%), High School or equivalent (28.25%), Master’s and above (24.16%) and no respondents reported “less than high school”. The most common response for online shopping behavior was “Monthly” accounting for 47.96% of the sample. Notably, the intentional AI usage variable revealed that 26.39% of respondents claim to use AI daily and 39.41% use it weekly; a much higher frequency than initially anticipated by the author. The implications and limitations of this will be discussed in Section 5.2 Limitations. Furthermore, the table below presents a summary of the collected data, while Appendix B contains the raw data tables.

Table 1. Summary of descriptive statistics (gender, age group, education, online shopping frequency, intentional AI usage frequency, initial preference, new preference and effective preference reversal)

Variable	Obs.	Mean	Std. dev.	Min.	Max.
gender	269	.539	.528	0	2
AgeG	269	1.13	.849	0	5
Educ	269	1.96	.724	1	3
OsFreq	269	1.71	.822	0	4
AuFreq	269	1.18	.958	0	4
INP	269	3.83	2.21	1	7
NWP	269	2.65	1.97	0	7
ERP	269	1.76	1.47	0	5

4.2 Testing assumptions & randomization

Before conducting the two-way ANOVA and t-tests, assumptions must be checked to ensure the validity and reliability of the results. To do this, the author tested for normality and homogeneity of variances. Furthermore, as the online experiment incorporated random assignment, the author tested if the randomization process was successful. Detailed tables and figures can be found in Appendix B.

4.2.1 Normality

To test for normality, the *EPR* variable was visually examined using a Q-Q plot and the Shapiro-Wilk W test. The Shipiro-Wilk W test revealed that the *EPR* variable does not follow a normal distribution, $W(269) = .975$, $p = .0001$, which is in accordance with the interpretation of the Q-Q Plot (Figure 8b.1 in Appendix B), hence not complying with the normality assumption. Despite this outcome, the two-way ANOVA and two-sample t-test are robust to certain degrees of non-normal data given that the homogeneity of variances assumption holds, and the sample size is sufficiently large ($n > 100$) (with equal distribution in test conditions). The one-sample t-test is also strong against deviations from normality when the sample size is large. However, the author used a Wilcoxon-signed rank test to confirm the t-test results, which is robust to non-normal data.

4.2.2 Homogeneity of Variances

To test the homogeneity of variances, Levene's test for equal variances using the *EPR* variable by the recommender and product type. The results for the recommender types were $F(1, 267) = 1.46$, $p = .228$, while the results for the product types were $F(1, 267) = 1.69$, $p = .194$, indicating that the assumption holds.

4.2.3 Randomization

To test if the random assignment was successful, a Chi-squared test was used on the scenario variable (representing the different test conditions) and genders of participants. The results were $\chi^2(6, N = 269) = 2.49$, $p = .869$, showing that the test conditions and genders had no significant relationship, thus confirming successful random assignment for the test conditions.

4.3 Hypothesis Testing

With the assumptions checked and the limitations in mind, the hypotheses can now be tested. The primary hypothesis: consumers will reverse their preferences after being exposed to a recommendation by either recommender type, was tested using a one-sample t-test and a Wilcoxon-signed rank test. This section will discuss the results of the statistical tests, and the detailed outputs from Stata can be found in Appendix B.

4.3.1 Hypothesis 1

The t-test revealed that the mean of the *EPR* was significantly different from zero, $t(268) = 19.8$, $p < .001$. This signifies that, on average, the recommenders were successful in reversing preferences with a mean *EPR* of 1.76 units (Min. = 0, Max. = 5). However, as mentioned, the *EPR* variable does not follow a normal distribution, which may invalidate the results of a t-test if sample sizes are not sufficiently large. Despite the sample size equaling 270 observations, to add confidence in the conclusion of the t-test, a Wilcoxon-signed rank test was used to test the median. The results of this test were $Z = 13.2$, $p < .001$, indicating that the median *EPR* is also non-zero.

With both the t-test and Wilcoxon-signed rank test in accordance, there is now strong evidence that EPR is significantly different from zero. This allows hypothesis 1 to be accepted, concluding that the recommenders were effective in inducing a preference reversal across all test conditions.

4.3.2 Hypotheses 2 and 3

To test hypotheses 2 and 3, a two-way ANOVA and a two-sample t-test were used. These tests revealed the main and interaction effects, as well as the recommender types' effectiveness in causing preference reversal changes for the different product types.

The two-way ANOVA results showed a significant interaction effect between recommender and product types on *EPR*, $F(1, 265) = 8.37$, $p = .0041$. This shows that the effectiveness of recommender types in causing a preference reversal is dependent on the product type. The main effect of recommender types was not significant, $F(1, 265) = .88$, $p = .3502$, showing that the recommender average *EPR* was similar when ignoring the product type. Similarly, the product type also does not cause significant changes in *EPR* when ignoring the recommender type; $F(1, 265) = .59$, $p =.4445$. This indicates that a particular recommender is only more effective than its counterpart for a specific product type. Furthermore, the overall two-way ANOVA model was significant ($F(3, 265) = 3.27$, $p = 0.0217$), but its $R^2 = .0357$ represents the low explanatory power of the model, meaning that the observed independent variables only accounted for 3.57% of the variance in the *EPR* variable. Section 5.2 Limitations will further discuss the limitations of this low explanatory power will be discussed further in Section 5.2 Limitations.

To get a deeper insight into the moderating effects of the product type, a two-sample t-test was used. The first two-sample t-test examined how the recommender types affected *EPR* when products were hedonic, which resulted in a significant difference in *EPR*, $t(132) = 2.644$, $p = 0.0092$. Thus, on average, the human recommender's *EPR* was approximately 49.4% (= 0.671 units) higher compared to AI recommenders when the product type is hedonic. This result allows the second hypothesis to be accepted, showing that human recommenders are more effective in causing preference reversals when considering hedonic products.

However, when examining the second two-sample t-test for when products were utilitarian, no significant difference in *EPR* between the recommender types was found, $t(133) = -1.4179$, *p*

= .1586. This result signifies that the AI recommenders were, on average, not more effective than human recommenders in causing preference reversals when considering utilitarian products. This shows that hypothesis 3 can be rejected

5 Discussion

5.1 Findings

As discussed by Donkers (2013) and shown in the previous section, an individual's preferences can be altered by changes that happen within their decision-making environment. To observe the extent of preference reversal, the author manipulated this environment by altering the recommender and product types.

The first hypothesis tested if recommenders (AI or human) were able to illicit a "preference reversal", a phenomenon that occurs when a person alters their preferences between a set of alternatives. Analyzing the variable *EPR* (effective preference reversal) post-recommender intervention allows for the calculation of the extent to which preferences can be reversed in a controlled environment. The results of the experiment showed that the recommenders were able to cause a change in preferences in 70.26% of observations across four test conditions, with an average *EPR* of 1.76 units (out of possible 6).

A deeper analysis was done by investigating the moderating effects of product type (hedonic or utilitarian) on *EPR*. Where hedonic goods are associated with enjoyment and self-fulfillment, utilitarian goods are commonly regarded as practical tools meant for a specific purpose. Past studies have examined the differences between these goods and how recommender types affect preferences toward them. They found that decisions regarding utilitarian goods are significantly more influenced by AI recommenders and a similar relationship exists between hedonic goods and human recommenders (Wien & Peluso, 2021). This paper pushed this phenomenon by investigating if the same conclusions can be drawn for a preference reversal scenario. Hypotheses 2 and 3 specifically questioned which recommender type (human vs. AI) is more effective at reversing preferences for each product type (hedonic vs. utilitarian)?

The findings showed that the recommender and product type variables, when independent of each other, were not able to cause significant differences in the effectiveness of reversing preferences. Thus, neither AI nor humans proved to be objectively more effective in reversing preference in all scenarios. Similarly, neither preference nor product type was swayed more easily than the other. However, accounting for both recommender and product types together revealed a highly significant effect, suggesting that a specific recommender is, on average, more effective at causing preference reversals for a specific product type.

A deeper analysis using a two-sample t-test revealed that human recommenders are, on average, significantly more effective at inducing a preference reversal when the product is hedonic. Longoni and Cian (2020) found that the close relationship between human characteristics and hedonic goals enables people to perceive humans as more competent recommending agents for hedonic-based decisions, compared to their AI counterparts. However, with the data gathered through the online experiment, the same conclusion cannot be drawn for AI recommenders. The results showed no significant difference in effectiveness to cause a preference reversal between the recommender types when considering utilitarian goods. This finding is contrary to the "word-of-machine" effect hypothesized by Longoni and Cian (2020) suggesting that algorithmic advice would be preferred over human advice for utilitarian-based decision. This result is more closely related to the "algorithmic aversion" discussed by Dietvorst et al. (2015), which describes how people often neglect the advice from machines despite their superior ability to make calculated judgements based on analyzing data.

5.2 Limitations

As discussed in the theoretical framework, consumer behavior is a highly complex field involving numerous factors that could influence the decision-making process. Covering all the topics and factors within consumer behavior is beyond the scope of the thesis. However, the author acknowledges the limitations of this experiment and provides recommendations for future research to achieve more reliable results.

5.2.1 Experiment Design Limitations

One such limitation is how the experiment uses stated preferences as a proxy for real preferences. This design choice creates a limitation as we now operate under the assumption that the “stated preferences” of the participants are equal to their “real preferences” (A.K.A. revealed preferences). In reality, the two may be different for various reasons, such as hypothetical bias, which occurs when people make choices that are not realistic due to them being in an experiment setting (Ajzen et al., 2004; Quaife et al., 2018). This can be considered a lesser limitation as, although not ideal, it is common practice as collecting data on stated preferences is much less resource intensive (De Corte et al., 2021).

A larger limitation of the experiment design lies in the intricacies of consumer behaviour. Numerous internal and external factors, some of which are typically unobservable, influence the formation of preferences and decision-making processes in an individual (Teleaba et al., 2021). It was not possible to account for a large number of variables within the confines of this paper. The consequences of limitation are especially apparent in the two-way ANOVA model used to examine the interaction effects of recommender and product types on preference reversals. While the overall model and interaction effects were found to be significant, the independent variables were only responsible for approximately 3.57 percent of the variance in *EPR*. This suggests that the independent variables may be weak predictors for *EPR*, or that other unmeasured variables play a larger role in reversing preferences.

Another notable limitation is the use of a Likert scale, which interprets the obtained data as interval variables despite its ordinal nature. The Likert’s scale is designed to rank items on an ordinal scale, which could lead to problems if it is interpreted as a continuous variable (Sullivan & Artino, 2013). A specific problem that was encountered was the non-normal distribution of the *EPR* variable, which may have been a consequence of the issues mentioned. However, the use of the Likert scale’s observations in this manner is not uncommon, where the problems with the ordinal nature can be somewhat mitigated when using more options in the scale (Sullivan & Artino, 2013).

5.2.2 Sampling Limitations

The experiment's results have limited external validity because of the chosen non-probability sampling methods. This was observed in the variable of “Intentional AI Use” (*AuFreq*), as the average 39.4% of individuals reported to be using AI on a weekly bias and 26.4% daily. This frequency of AI usage was higher than anticipated and cannot be easily compared to other datasets because AI is still a relatively new topic. A study by Kennedy et al. (2023), explored the how often Americans recognized interaction with AI, with 27% reporting “a few times a day” and another 28% reporting “a few times a week”. Although the frequency is similar to what was found from this research, the variable of “recognizing interaction with AI” and “intentional usage of AI” have differences in the aspect of intention, thus not entirely justifying the high frequency observed.

One plausible explanation for the result is that the majority of respondents were active university students at the release of ChatGPT in November 2022. During this period, the use of AI tools was a frequently discussed topic in academia, which likely led to increased exposure and familiarity, ultimately resulting in more frequent use of AI systems than other age groups. Another cause may be the method of distributing the experiment. The distribution was through social media platforms (using convenience and snowball sampling methods) and revealed the topic of the experiment as involving artificial intelligence. As participation was entirely voluntary, this could have resulted in the overrepresentation of individuals familiar with AI.

5.3 Future Research

The avenues for future research on the relationship between recommender types, product types, and preference reversals are vast. Especially since the use of AI is becoming exponentially popular, there remain many questions unanswered and variables unobserved.

Future research could begin by examining variables commonly associated with decision-making, which this paper did not cover. Some interesting suggestions include investigating the potential effects of individual expertise on a product, trust in AI, prior experiences with recommenders (helpful or harmful), and products with different decision involvement levels. All the variables listed could provide interesting insights into the effectiveness of recommenders under different conditions, allowing for a deeper understanding of which factors play larger roles in potentially reversing preferences.

Another aspect that the author recommends exploring are the future generations, specifically looking at how kids who are born in 2020 and beyond will interact with AI in the future. Whereas the current generation of young adults is often described as having an affinity for the internet, those who are born today may have a similar associated with AI, making them ideal candidates to observe in the context of AI vs. human recommenders. Furthermore, by conducting a similar experiment in the future, the advancement of AI systems is guaranteed, and their overall effectiveness would be greatly improved. Perhaps exploring the differences between a digital chatbot and a physical robot would also give new insights.

Lastly, to get results that more closely resemble revealed preferences, instead of the stated preferences used in this paper, an experiment involving the realization of the consequences of participant choices may be interesting. A recommended method to do this is to "simply" award the participant the product that they most prefer. This creates a real consequence for the participant's choices and allows observation of real preferences as the participants now have something to gain based on their decision, which may give a more accurate results in a recommender's effectiveness to reverse preferences.

6 Conclusion

This thesis aimed to investigate the relationship between recommender types and their effectiveness in causing a preference reversal (i.e. changing one's opinion), while also exploring the moderating effects of product types.

This was done by examining the following hypotheses:

H1. *Consumers will reverse their preferences after being exposed to a recommendation by either recommender type*

H2. *Human recommenders cause significantly higher EPR than AI recommenders when considering hedonic products*

H3. *AI recommenders cause significantly higher EPR than human recommenders when considering utilitarian products*

The data used to answer the hypotheses was gathered through an online experiment, distributed through social media platforms, and targeted young adults between 18-29 years old. The experiment was made up of four randomly-assigned test conditions, allocating participants with a hedonic or utilitarian and a human or AI recommender attempted to change their preferences. By capturing the participant's preference before and after recommender intervention, the variable of effective preference reversal *(EPR)* was calculated to measure the effectiveness of the recommenders in reversing preferences. On average, across all conditions, both recommenders were successful in reversing preferences to a significant degree, thus accepting the first hypothesis.

Furthermore, examining the moderating effects of product type revealed that for hedonic products, human recommenders were significantly more effective in reversing preferences, with an average *EPR* 49.4% higher than the AI alternative. This finding was in line with the second hypothesis and suggests a strong relationship between hedonic products and human advice, which may be due to the similarities of attributes associated with them. Contrastingly, the AI recommenders had no significant advantage compared to human recommenders in terms of their effectiveness to reverse preferences for utilitarian products. This outcome may have been caused by the human tendency to avoid advice received from machines, also known as algorithmic aversion.

Thus, answering the question: are AI recommenders more effective than their human counterparts in reversing preferences? This research concluded that humans are equally capable, and in certain conditions, they are even more effective than AI recommenders in persuading consumers to change their preferences.

The findings of this paper show that investment in a recommender agent (whether human or AI) can help businesses win over more customers despite increasing competition in the marketplace. New entrants to competitive markets would especially benefit from recommending systems that directly compare their products versus competitors in the same space. To maximize the effectiveness of the recommenders, businesses should explore what their product inherently represents for their target audience – a hedonic pursuit or a utilitarian one. Regarding the application of AI, although this experiment showed no definite evidence that AI recommenders perform better than their human counterparts, the inherent nature of AI allow for constant improve as more data is made available to. An early investment in an AI recommendation system may still be a highly beneficial endeavour in the long run, despite the lack of evidence today.

7 References

Ajzen, I., Brown, T. C., & Carvajal, F. (2004). Explaining the Discrepancy between Intentions and Actions: The Case of Hypothetical Bias in Contingent Valuation. *Personality & Social Psychology Bulletin*, *30*(9), 1108–1121. https://doi.org/10.1177/0146167204264079

Allen, W. (1953). The influence of source credibility on communication effectiveness. *Educational Technology Research and Development*, *1*(2), 142–143. https://doi.org/10.1007/bf02716996

Appel, G., Grewal, L., Hadi, R., & Stephen, A. T. (2019). The future of social media in marketing. *Journal of the Academy of Marketing Science*, *48*(1), 79–95. https://doi.org/10.1007/s11747-019-00695-1

Araujo, T. (2018). Living up to the chatbot hype: The influence of anthropomorphic design cues and communicative agency framing on conversational agent and company perceptions. *Computers in Human Behavior*, *85*, 183–189. https://doi.org/10.1016/j.chb.2018.03.051

Ariely, D. (2010). *Predictably irrational, revised and expanded edition: The Hidden Forces That Shape Our Decisions*. Harper Collins.

Batra, R., & Ahtola, O. T. (1991). Measuring the hedonic and utilitarian sources of consumer attitudes. *Marketing Letters*, *2*(2), 159–170. https://doi.org/10.1007/bf00436035

Boerman, S. C., Kruikemeier, S., & Borgesius, F. J. Z. (2017). Online Behavioral Advertising: A Literature review and Research agenda. *Journal of Advertising*, *46*(3), 363–376. https://doi.org/10.1080/00913367.2017.1339368

Botti, S., & McGill, A. L. (2011). The Locus of Choice: Personal Causality and Satisfaction with Hedonic and Utilitarian Decisions. ˜ *the œ Journal of Consumer Research/Journal of Consumer Research*, *37*(6), 1065–1078. https://doi.org/10.1086/656570

Budzinski, O., Gaenssle, S., & Lindstädt, N. (2021). Data (R)Evolution – The economics of algorithmic search and recommender services. *Social Science Research Network*. https://doi.org/10.2139/ssrn.3769386

Cambridge Business English Dictionary. (2024). Consumer preference. In *Cambridge Dictionary*. https://dictionary.cambridge.org/dictionary/english/consumer-preference

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

Çelen, B., Kariv, S., & Schotter, A. (2010). An experimental test of advice and social learning. *Management Science*, *56*(10), 1687–1701. https://doi.org/10.1287/mnsc.1100.1228

Constant, D., Sproull, L., & Kiesler, S. (1996). The kindness of strangers: the usefulness of electronic weak ties for technical advice. *Organization Science*, *7*(2), 119–135. https://doi.org/10.1287/orsc.7.2.119

Dawes, R. M. (1979). The robust beauty of improper linear models in decision making. *American Psychologist/˜ the œ American Psychologist*, *34*(7), 571–582. https://doi.org/10.1037/0003-066x.34.7.571

De Bruyn, A., & Lilien, G. L. (2008). A multi-stage model of word-of-mouth influence through viral marketing. *International Journal of Research in Marketing*, *25*(3), 151–163. https://doi.org/10.1016/j.ijresmar.2008.03.004

De Corte, K., Cairns, J., & Grieve, R. (2021). Stated versus revealed preferences: An approach to reduce bias. *Health Economics*, *30*(5), 1095–1123. https://doi.org/10.1002/hec.4246

Dietvorst, B. J., Simmons, J. P., & Massey, C. (2015). Algorithm aversion: People erroneously avoid algorithms after seeing them err. *Journal of Experimental Psychology. General*, *144*(1), 114–126. https://doi.org/10.1037/xge0000033

Donkers, B. (2013). *The Customer Cannot Choose*. Erasmus Research Institute of Management.

Edwards, W. (1954). The theory of decision making. *Psychological Bulletin*, *51*(4), 380–417. https://doi.org/10.1037/h0053870

Garvey, A. M., Meloy, M. G., & Shiv, B. (2017). The jilting Effect: antecedents, mechanisms, and consequences for preference. *Journal of Marketing Research*, *54*(5), 785–798. https://doi.org/10.1509/jmr.14.0373

Gilovich, T., Griffin, D., & Kahneman, D. (2002). Heuristics and biases. In *Cambridge University Press eBooks*. https://doi.org/10.1017/cbo9780511808098

Grove, W. M., & Meehl, P. E. (1996). Comparative efficiency of informal (subjective, impressionistic) and formal (mechanical, algorithmic) prediction procedures: The clinical–statistical controversy. *Psychology, Public Policy, and Law*, *2*(2), 293–323. https://doi.org/10.1037/1076-8971.2.2.293

Hansson, S. O. (1995). Changes in preference. *Theory and Decision*, *38*(1), 1–28. https://doi.org/10.1007/bf01083166

Häubl, G., & Murray, K. B. (2003). Preference Construction and Persistence in Digital Marketplaces: The Role of Electronic Recommendation Agents. *Journal of Consumer Psychology*, *13*(1), 75–91. https://doi.org/10.1207/153276603768344807

Hennig-Thurau, T., Gwinner, K. P., Walsh, G., & Gremler, D. D. (2004). Electronic word-of-mouth via consumer-opinion platforms: What motivates consumers to articulate themselves on the Internet? *Journal of Interactive Marketing*, *18*(1), 38–52. https://doi.org/10.1002/dir.10073

Hoyer, W. D., MacInnis, D. J., Pieters, R., Chan, E., & Northey, G. (2017). *Consumer Behaviour: Asia-Pacific Edition*. Cengage AU.

Jin, F., & Zhang, X. (2023). Artificial intelligence or human: when and why consumers prefer AI recommendations. *Information Technology & People*. https://doi.org/10.1108/itp-01-2023-0022

Johnson, E. J., Payne, J. W., & Bettman, J. R. (1988). Information displays and preference reversals. *Organizational Behavior and Human Decision Processes*, *42*(1), 1–21. https://doi.org/10.1016/0749-5978(88)90017-9

Kennedy, B., Tyson, A., & Saks, E. (2023, February 15). *Public awareness of artificial intelligence in everyday activities*. Pew Research Center. Retrieved August 14, 2024, from https://www.pewresearch.org/science/2023/02/15/public-awareness-of-artificial-intelligence-in-everyday-activities/

Kim, J., Giroux, M., & Lee, J. C. (2021). When do you trust AI? The effect of number presentation detail on consumer trust and acceptance of AI recommendations. *Psychology & Marketing*, *38*(7), 1140–1155. https://doi.org/10.1002/mar.21498

Kim, T., & Duhachek, A. (2023). Artificial intelligence and consumer Psychology. In *Cambridge University Press eBooks* (pp. 247–270). https://doi.org/10.1017/9781009243957.011

Kim, T. W., & Duhachek, A. (2020). Artificial Intelligence and Persuasion: A Construal-Level account. *Psychological Science*, *31*(4), 363–380. https://doi.org/10.1177/0956797620904985

Lee, L., Lee, M. P., Bertini, M., Zauberman, G., & Ariely, D. (2015). Money, time, and the stability of consumer preferences. *Journal of Marketing Research*, *52*(2), 184–199. https://doi.org/10.1509/jmr.10.0386

Levin, J., & Milgrom, P. (2004). *Introduction to Choice Theory* [PDF].

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

Lu, J., Liu, Z., & Fang, Z. (2016). Hedonic products for you, utilitarian products for me. *Judgment and Decision Making*, *11*(4), 332–341. https://doi.org/10.1017/s1930297500003764

McKenzie, C. M. R. (2018). Constructed preferences, rationality, and choice architecture. *Review of Behavioral Economics*, *5*(3–4), 337–370. https://doi.org/10.1561/105.00000091

Mende, M., Scott, M. L., Van Doorn, J., Grewal, D., & Shanks, I. (2019). Service Robots Rising: How humanoid robots influence service experiences and elicit compensatory consumer responses. *Journal of Marketing Research*, *56*(4), 535–556. https://doi.org/10.1177/0022243718822827

Moustakas, E., Lamba, N., Mahmoud, D., & Ranganathan, C. (2020). Blurring lines between fiction and reality: Perspectives of experts on marketing effectiveness of virtual influencers. *2020 International Conference on Cyber Security and Protection of Digital Services (Cyber Security)*. https://doi.org/10.1109/cybersecurity49315.2020.9138861

Önkal, D., Goodwin, P., Thomson, M., Gönül, S., & Pollock, A. (2009). The relative influence of advice from human experts and statistical methods on forecast adjustments. *Journal of Behavioral Decision Making*, *22*(4), 390–409. https://doi.org/10.1002/bdm.637

Ozdemir, O., Kolfal, B., Messinger, P. R., & Rizvi, S. (2023). Human or virtual: How influencer type shapes brand attitudes. *Computers in Human Behavior*, *145*, 107771. https://doi.org/10.1016/j.chb.2023.107771

Quaife, M., Terris-Prestholt, F., Di Tanna, G. L., & Vickerman, P. (2018). How well do discrete choice experiments predict health choices? A systematic review and meta-analysis of external validity. *The European Journal of Health Economics*, *19*(8), 1053–1066. https://doi.org/10.1007/s10198-018-0954-6

Regenwetter, M., Dana, J., & Davis-Stober, C. P. (2011). Transitivity of preferences. *Psychological Review*, *118*(1), 42–56. https://doi.org/10.1037/a0021150

Sah, S., & Loewenstein, G. (2014). Nothing to declare. *Psychological Science*, *25*(2), 575–584. https://doi.org/10.1177/0956797613511824

Sanders, N. R., & Manrodt, K. B. (2003). The efficacy of using judgmental versus quantitative forecasting methods in practice. *Omega*, *31*(6), 511–522. https://doi.org/10.1016/j.omega.2003.08.007

Schecter, A., Bogert, E., & Lauharatanahirun, N. (2023). Algorithmic appreciation or aversion? The moderating effects of uncertainty on algorithmic decision making. *CHI EA '23: Extended Abstracts of the 2023 CHI Conference on Human Factors in Computing Systems*. https://doi.org/10.1145/3544549.3585908

Schiffman, L. G., & Wisenblit, J. L. (2018). *Consumer Behavior, Global Edition*. Pearson Higher Ed.

Senecal, S., & Nantel, J. (2004). The influence of online product recommendations on consumers' online choices. *Journal of Retailing*, *80*(2), 159–169. https://doi.org/10.1016/j.jretai.2004.04.001

Shaffer, V. A., Probst, C. A., Merkle, E. C., Arkes, H. R., & Medow, M. A. (2012). Why do patients derogate physicians who use a Computer-Based Diagnostic Support System? *Medical Decision Making*, *33*(1), 108–118. https://doi.org/10.1177/0272989x12453501

Sher, S., & McKenzie, C. R. M. (2014). Options as information: Rational reversals of evaluation and preference. *Journal of Experimental Psychology General*, *143*(3), 1127–1143. https://doi.org/10.1037/a0035128

Slovic, P. (1995). The construction of preference. *American Psychologist/˜ the œAmerican Psychologist*, *50*(5), 364–371. https://doi.org/10.1037/0003-066x.50.5.364

Strohmaier, D., & Messerli, M. (2024). *Preference change*. https://doi.org/10.1017/9781009181860

Sullivan, G. M., & Artino, A. R. (2013). Analyzing and interpreting data from Likert-Type scales. *Journal of Graduate Medical Education*, *5*(4), 541–542. https://doi.org/10.4300/jgme-5-4-18

Teleaba, F., Popescu, S., Olaru, M., & Pitic, D. (2021). Risks of observable and unobservable biases in artificial intelligence predicting consumer Choice. *Amfiteatru Economic*, *23*(56), 102. https://doi.org/10.24818/ea/2021/56/102

Ulen, T. (1999). *Encyclopedia of Law and Economics: The history and methodology of law and economics* (B. Bouckaert & G. De Geest, Eds.; 1st ed., Vol. 0710). Edward Elgar.

Wang, X., Lu, Z., & Yin, M. (2022). Will you accept the AI recommendation? Predicting Human Behavior in AI-Assisted Decision Making. *Proceedings of the ACM Web Conference 2022*. https://doi.org/10.1145/3485447.3512240

Wardman, M. (1988). A COMPARISON OF REVEALED PREFERENCE AND STATED PREFERENCE MODELS OF TRAVEL BEHAVIOUR. *Journal of Transport Economics and Policy*, *22*(1). https://trid.trb.org/view/289552

Whitson, D. A., & Henry, W. A. (1996). What's in a label? Environmental issues in product packaging. *Journal of Euromarketing/Journal of Euro-marketing*, *5*(3), 29–42. https://doi.org/10.1300/j037v05n03_03

Whitson, D., Ozkaya, H. E., & Roxas, J. (2014). Changes in consumer segments and preferences to green labelling. *International Journal of Consumer Studies*, *38*(5), 458–466. https://doi.org/10.1111/ijcs.12103

Wien, A. H., & Peluso, A. M. (2021). Influence of human versus AI recommenders: The roles of product type and cognitive processes. *Journal of Business Research*, *137*, 13–27. https://doi.org/10.1016/j.jbusres.2021.08.016

8 Appendix

Appendix A: Online Experiment

Item 8a.1 Introduction

Dear Participant,

Thank you for taking the time to participate in our study.
This experiment will take approximately. 2 minutes to complete.

We are investigating how different types of recommenders and product types influence preferences.
Your participation will involve answering a few questions about your preferences in hypothetical circumstances.

Please note that your identity and responses will be kept completely confidential, and the data collected will be used solely for research purposes. Your participation is voluntary, and you may withdraw from the survey at any time without any consequences.

If you agree with the terms and consent to your answers being used for a study please select "Yes" to continue.

Item 8a.2 Basic Demographics

Question	Answering options
What best describes you?	Male / Female / Non-binary / Prefer not to say
Which age group do you currently fall into?	18-21 / 22-25 / 26-29 / 30-33 / 34-37 / 38+
What is the highest level of education you have achieved?	Less than High School / High School or equivalent / Bachelor's Degree / Master's Degree or higher / Prefer not to say
What best describes your online shopping frequency?	Daily / Weekly / Monthly / Yearly / Never
What best described your habits on intentional AI technology use? Examples: Chatbots (ChatGPT), virtual assistants (Siri), image generators (GPT-4, DALL-E), etc.	Daily / Weekly / Monthly / Yearly / Never

Item 8a.3 Hedonic Product Selection

You are shopping for a pair of headphones. You plan to use to use the headphones for leisure activities (listening to music, podcasts, watching videos, etc.). It is important that these headphones give you an enjoyable experience.

You are searching online for the best headphones that fit your needs and have narrowed the choice down to the two following products:

	Product A	**Product B**
Name	EchoSound 4	SonixPro Max
Connectivity	Wireless only	Wireless & wired
Battery Life	6h - 8h	6h (Wireless)
Sound Profile	Bass Heavy	Adjustable
Noise Cancelling	Yes, Active Noise Cancelling	Yes, Active Noise Cancelling

Question	Answering options
How would you describe your preference towards the products above?	Strongly Prefer Product A / Prefer Product A / Slightly Prefer Product A / Indifferent or Neutral / Slightly Prefer Product B / Strongly Prefer Product B

Item 8a.4 Utilitarian Product Selection

You are shopping for a new laptop. You plan to use the laptop for work/school (writing documents, attending meetings, making presentations, etc.)

It is important that the laptop be efficient in helping you complete work/school-related tasks.

You are searching online for the best laptop that fits your needs and have narrowed the choice down to the two following products:

	Product A	**Product B**
Name	TechBook Slim	AeroTop Lite
Storage	1TB SSD	500GB SSD, 1TB HDD
Battery Life	16h - 20h	14h - 18h
Weight	1.5kg	1.9kg
Screen	13.6in, 4K Display	15in, 1440p Display

Question	**Answering options**
How would you describe your preference towards the products above?	Strongly Prefer Product A / Prefer Product A / Slightly Prefer Product A / Indifferent or Neutral / Slightly Prefer Product B / Strongly Prefer Product B

Item 8a.5 Human Recommendations (4 possible recommendations)

You showed a preference towards (Product A / Product B) (To be indifferent/neutral).

Following your initial choice, you find a recommendation left by a previous customer who shares their insights on the products you were comparing.

Customer Recommendation:

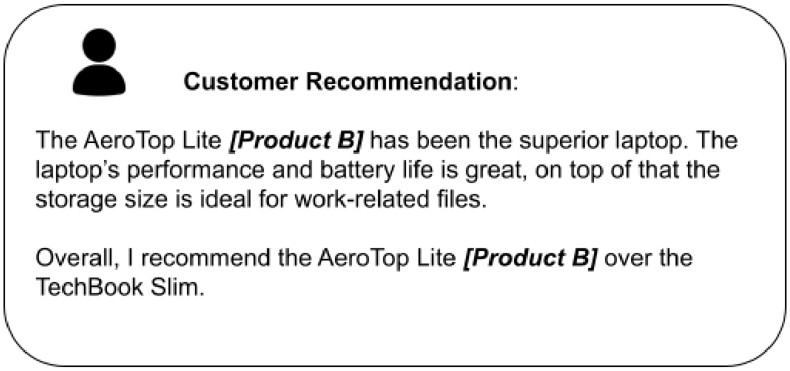
The SonixPro Max ***[Product B]*** has been the superior set of headphones. Sound quality is great, noise cancelling has been fantastic and the battery life is great for on-the-go listening.

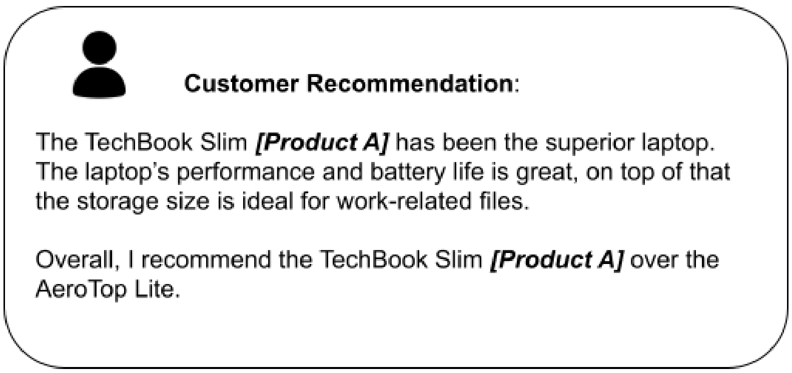
Overall, I recommend the SonixPro Max ***[Product B]*** over the EchoSound 4.

Customer Recommendation:

The EchoSound 4 ***[Product A]*** has been the superior set of headphones. Sound quality is great, noise cancelling has been fantastic and the battery life is great for on-the-go listening.

Overall, I recommend the EchoSound 4 ***[Product A]*** over the SonixPro Max.





Question	**Answering options**
After seeing the recommendation, has your preference changed?	Yes / No
Please specify how you would describe your **new preference**?	Strongly Prefer Product A / Prefer Product A / Slightly Prefer Product A / Indifferent or Neutral / Slightly Prefer Product B / Strongly Prefer Product B

Item 8a.6 AI Recommendations (4 possible recommendations)

You showed a preference towards (Product A / Product B) (To be indifferent/neutral).

Following your previous choice, you find a recommendation by an artificial intelligence (AI) recommender who shares insights on the products you were comparing.

Note that this AI was trained using previous customer data.

AI Recommendation:

The SonixPro Max ***[Product B]*** has been the superior set of headphones. Sound quality is great, noise cancelling has been fantastic and the battery life is great for on-the-go listening.

Overall, I recommend the SonixPro Max ***[Product B]*** over the EchoSound 4.

AI Recommendation:

The EchoSound 4 ***[Product A]*** has been the superior set of headphones. Sound quality is great, noise cancelling has been fantastic and the battery life is great for on-the-go listening.

Overall, I recommend the EchoSound 4 ***[Product A]*** over the SonixPro Max.

AI Recommendation:

The AeroTop Lite ***[Product B]*** has been the superior laptop. The laptop's performance and battery life is great, on top of that the storage size is ideal for work-related files.

Overall, I recommend the AeroTop Lite ***[Product B]*** over the TechBook Slim.

AI Recommendation:

The TechBook Slim ***[Product A]*** has been the superior laptop. The laptop's performance and battery life is great, on top of that the storage size is ideal for work-related files.

Overall, I recommend the TechBook Slim ***[Product A]*** over the AeroTop Lite.

Question	**Answering options**
After seeing the recommendation, has your preference changed?	Yes / No
Please specify how you would describe your **new preference**?	Strongly Prefer Product A / Prefer Product A / Slightly Prefer Product A / Indifferent or Neutral / Slightly Prefer Product B / Strongly Prefer Product B

Appendix B: Tables and Figures

Table 8b.1 Descriptive Statistics of Gender

gender	Freq.	Percent	Cum.
Female	128	47.58	47.58
Male	137	50.93	98.51
Other	4	1.49	100.00
Total	269	100.00	

Table 8b.2 Descriptive Statistics of Age Group

ageG	Freq.	Percent	Cum.
18-21	49	18.22	18.22
22-25	158	58.74	76.95
26-29	46	17.10	94.05
30-33	11	4.09	98.14
34-37	3	1.12	99.26
38+	2	0.74	100.00
Total	269	100.00	

Table 8b.3 Descriptive Statistics of Education

educ	Freq.	Percent	Cum.
HS/Equivalent	76	28.25	28.25
Bachelor's	128	47.58	75.84
Master's +	65	24.16	100.00
Total	269	100.00	

Table 8b.4 Descriptive Statistics of Online Shopping Frequency

OsFreq	Freq.	Percent	Cum.
Daily	10	3.72	3.72
Weekly	99	36.80	40.52
Monthly	129	47.96	88.48
Yearly	20	7.43	95.91
Never	11	4.09	100.00
Total	269	100.00	

Table 8b.5 Descriptive Statistics of Intentional AI Usage Frequency

AuFreq	Freq.	Percent	Cum.
Daily	71	26.39	26.39
Weekly	106	39.41	65.80
Monthly	68	25.28	91.08
Yearly	20	7.43	98.51
Never	4	1.49	100.00
Total	269	100.00	

Table 8b.6 Descriptive Statistics of Scenarios (0 = hedonic product & human recommender, 1 = hedonic product & AI recommender, 2 = utilitarian product & human recommender, 3 = utilitarian product & AI recommender)

scen	Freq.	Percent	Cum.
0	67	24.91	24.91
1	67	24.91	49.81
2	67	24.91	74.72
3	68	25.28	100.00
Total	269	100.00	

Table 8b.7 Descriptive Statistics of INP

INP	Freq.	Percent	Cum.
Strng. A	52	19.33	19.33
Pref. A	56	20.82	40.15
Slght. A	27	10.04	50.19
IDF	18	6.69	56.88
Slght. B	19	7.06	63.94
Pref. B	61	22.68	86.62
Strng. B	36	13.38	100.00
Total	269	100.00	

Table 8b.8 Descriptive Statistics of change in preference; denotes if there was a change in preference after recommender intervention

Chng	Freq.	Percent	Cum.
No	80	29.74	29.74
Yes	189	70.26	100.00
Total	269	100.00	

Table 8b.8 Descriptive Statistics of EPR

EPR	Freq.	Percent	Cum.
0	80	29.74	29.74
1	38	14.13	43.87
2	58	21.56	65.43
3	57	21.19	86.62
4	31	11.52	98.14
5	5	1.86	100.00
Total	269	100.00	

Table 8b.9 Descriptive Statistics of initial preferences (INP) by scenario

scen	INP							
	Strng. A	Pref. A	Slght. A	IDF	Slght. B	Pref. B	Strng. B	Total
0	9	8	11	3	4	22	10	67
1	11	12	5	5	7	13	14	67
2	21	18	5	3	3	11	6	67
3	11	18	6	7	5	15	6	68
Total	52	56	27	18	19	61	36	269

Table 8b.10 Descriptive statistics of change in preference (Chng) by scenario; denotes if there was a change in preference after recommender intervention

scen	Chng		
	No	Yes	Total
0	18	49	67
1	25	42	67
2	22	45	67
3	15	53	68
Total	80	189	269

Table 8b.11 Descriptive statistics of EPR by scenario

scen	EPR						
	0	1	2	3	4	5	Total
0	18	9	13	10	14	3	67
1	25	13	12	14	3	0	67
2	22	8	15	15	7	0	67
3	15	8	18	18	7	2	68
Total	80	38	58	57	31	5	269

Table 8b.12 Shapiro-Wilk W test for normal data

Shapiro-Wilk W test for normal data

Variable	Obs	W	V	z	Prob>z
EPR	269	0.97487	4.861	3.692	0.00011

Table 8b.13 Levene's test for homogeneity of variance (recommender type)

RecType	Summary of EPR Mean	Std. dev.	Freq.
Human	1.8432836	1.5259878	134
AI	1.6814815	1.3857045	135
Total	1.7620818	1.4568075	269

W0 = 1.4578612 df(1, 267) Pr > F = 0.22833924

W50 = 1.3286141 df(1, 267) Pr > F = 0.25008284

W10 = 1.1568637 df(1, 267) Pr > F = 0.28308789

Table 8b.14 Levene's test for homogeneity of variance (product type)

ProdType	Summary of EPR Mean	Std. dev.	Freq.
Hedonic	1.6940299	1.5030176	134
Utilitari	1.8296296	1.411788	135
Total	1.7620818	1.4568075	269

W0 = 1.6937309 df(1, 267) Pr > F = 0.19423171

W50 = 2.1241909 df(1, 267) Pr > F = 0.14616445

W10 = 1.5001852 df(1, 267) Pr > F = 0.22172235

Table 8b.15 Chi-square test on scenario by gender; to check for successful random assignment

	gender			
scen	Female	Male	Other	Total
0	30	36	1	67
1	34	33	0	67
2	31	35	1	67
3	33	33	2	68
Total	128	137	4	269

Pearson chi2(6) = 2.4920 Pr = 0.869

Table 8b.16 One-sample t-test on EPR

One-sample t test

Variable	Obs	Mean	Std. err.	Std. dev.	[95% conf.	interval]
EPR	269	1.762082	.0888231	1.456808	1.587202	1.936962

mean = mean(EPR) t = 19.8381

H0: mean = 0 Degrees of freedom = 268

Ha: mean < 0	Ha: mean != 0	Ha: mean > 0
Pr(T < t) = 1.0000	Pr(|T| > |t|) = 0.0000	Pr(T > t) = 0.0000

Table 8b.17 Wilcoxon-signed rank test on EPR

Wilcoxon signed-rank test

Sign	Obs	Sum ranks	Expected
Positive	189	33075	16537.5
Negative	0	0	16537.5
Zero	80	3240	3240
All	269	36315	36315

Unadjusted variance	1631148.75
Adjustment for ties	-9685.50
Adjustment for zeros	-43470.00
Adjusted variance	1577993.25

H0: EPR = 0

z = 13.165

Prob > |z| = 0.0000

Table 8b.18 Two-Way ANOVA; *EPR* on recommender type and product type

```
Number of obs =        269     R-squared     =  0.0357
Root MSE      =    1.43862     Adj R-squared =  0.0248
```

Source	Partial SS	df	MS	F	Prob>F
Model	20.325473	3	6.7751577	3.27	0.0217
RecType	1.8126342	1	1.8126342	0.88	0.3502
ProdType	1.2134163	1	1.2134163	0.59	0.4445
RecType#ProdType	17.317398	1	17.317398	8.37	0.0041
Residual	548.44776	265	2.0696142		
Total	568.77323	268	2.1222882		

Table 8b.19 Two-sample t-test on *EPR*, product type = hedonic, by recommender type

Two-sample t test with equal variances

Group	Obs	Mean	Std. err.	Std. dev.	[95% conf.	interval]
Human	67	2.029851	.1983311	1.62341	1.63387	2.425831
AI	67	1.358209	.1587227	1.299201	1.041309	1.675109
Combined	134	1.69403	.1298409	1.503018	1.437209	1.95085
diff		.6716418	.2540238		.1691575	1.174126

```
    diff = mean(Human) - mean(AI)                                t =   2.6440
H0: diff = 0                                    Degrees of freedom =      132

    Ha: diff < 0                 Ha: diff != 0                 Ha: diff > 0
 Pr(T < t) = 0.9954         Pr(|T| > |t|) = 0.0092          Pr(T > t) = 0.0046
```

Table 8b.20 Two-sample t-test on *EPR*, product type = utilitarian, by recommender type

Two-sample t test with equal variances

Group	Obs	Mean	Std. err.	Std. dev.	[95% conf.	interval]
Human	67	1.656716	.1721866	1.409408	1.312935	2.000498
AI	68	2	.1702139	1.40362	1.660252	2.339748
Combined	135	1.82963	.1215074	1.411788	1.589309	2.06995
diff		-.3432836	.2421103		-.8221684	.1356013

```
    diff = mean(Human) - mean(AI)                                 t =  -1.4179
H0: diff = 0                                     Degrees of freedom =       133

    Ha: diff < 0                 Ha: diff != 0                 Ha: diff > 0
 Pr(T < t) = 0.0793         Pr(|T| > |t|) = 0.1586          Pr(T > t) = 0.9207
```

Figure 8b.1 Q-Q Plot *(EPR)*