STOCKHOLM SCHOOL OF ECONOMICS
1351 Master's Thesis in Business Management Academic year 2025–2026

How may (A)I assist you?

The Effect of Price Level and Transparency on User Intention in an AI Recommender Setting

Hamza El Fakir (24958) and Philip Schrader (42789)

Abstract:

AI-assisted booking tools are becoming increasingly common in online travel. When such systems recommend a flight, two questions arise: does explaining the reasoning behind the recommendation make consumers more willing to book, and does price shape this effect? This study examines these questions using a 2×2 between-subjects experiment with Swedish adults aged 20–39 (N = 364). Participants were randomly assigned to one of four conditions, with price level (high vs. low) and AI transparency (no explanation vs. a structured explanation outlining selection criteria and data sources) as factors. Purchase intention was the primary outcome, and trust and perceived financial risk were examined as potential mediators. The study did not find empirical support for price as a formal boundary condition, as the interaction between price and transparency on purchase intention was not significant. Contrast analyses, however, indicated that the positive effect of transparency on purchase intention appeared only in the high-price condition. At the same time, the transparency manipulation did not produce a meaningful difference in perceived transparency in the low-price condition, leaving the overall pattern inconclusive. Trust and perceived financial risk, which were identified as psychological mechanisms through which transparency was expected to influence purchase intention, did not mediate this relationship. For airlines and online travel agencies, this points to high-stakes contexts as a potentially more promising target for explainability features, while further research is needed to establish when and how transparency reliably affects consumer decisions.

Keywords:

AI transparency, Explainable AI, Recommender systems, Consumer trust, Perceived risk, Purchase Intention, Online Flight booking

Supervisor: Magnus Söderlund

Date examined: 29.05.2026

Discussant(s): Hoang Nguyen

Examiners: Lily Gao & Alexander Mafael

Acknowledgements

We wish to express our deepest gratitude to our supervisor, Magnus Söderlund, for his incredible guidance and thoughtful feedback throughout the whole writing process. His ideas and insights have significantly impacted this thesis.

We would also like to extend our greatest appreciation to Norstat for generously supporting our research through the pro bono distribution of our experimental study. Without their contribution, this study would not be possible. Special thanks to Emma, who was our main point of contact throughout the data collection process and ensured we obtained the highest-quality data possible.

Finally, we want to thank everyone who took the time to ensure an accurate Swedish translation of our study, as well as all those who participated in our survey.

Hamza & Philip

Table of Contents

List of Tables

List of Figures

1. Introduction

In 1964, Donald F. Cox and Stuart U. Rich published an article titled "Perceived Risk and Consumer Decision-Making: The Case of Telephone Shopping", which explored the then under-studied role of perceived risk in a consumer behaviour context. Focusing on the empirical context of telephone shopping, the authors studied the strategies individuals adopt to mitigate risk, which they defined as a "function of the amount at stake and the certainty of the outcome" (Cox & Rich, 1964). They found that while it is rarely possible to reduce the amount at stake, consumers can increase certainty through information seeking. Specifically, they identified two types of consumers: those who rely on experience-based information and those who seek external information, such as product advertisements.
Today, more than 50 years later, while telephone shopping is no longer a significant phenomenon, many parallels can be drawn to AI recommender systems (see chapter 1.2). One particularly intriguing parallel concerns the role of information as a risk-reduction mechanism: just as newspaper advertising provided additional information to increase certainty in Cox and Rich’s study, AI explainability and transparency in recommender systems provide users with information intended to reduce uncertainty about a specific recommendation.

1.1 The Role of Recommender Systems and the Customer Journey in the Flight Booking Context

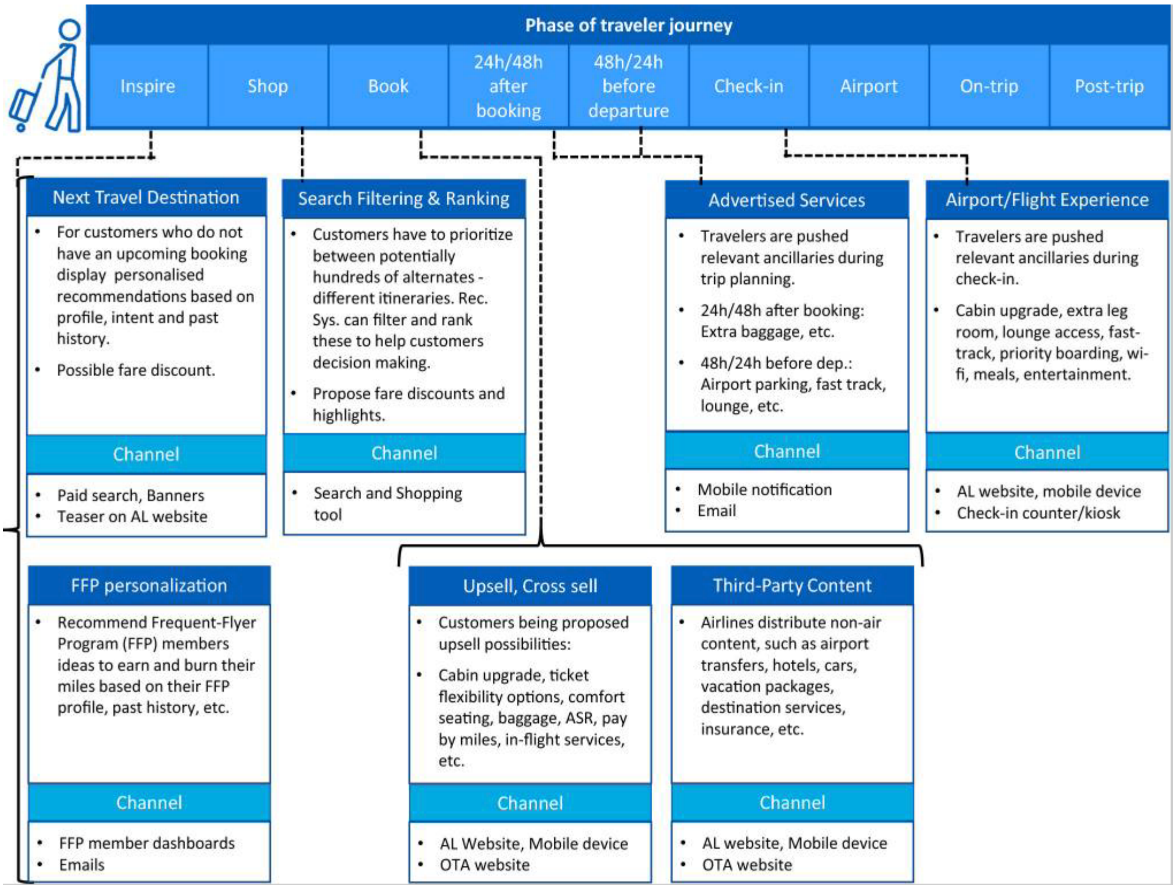
When booking a flight online, consumers face a large and often overwhelming number of alternatives, including different airlines, fare classes, connection types, and prices. Research shows that under these conditions, evaluating every option in depth triggers cognitive overload, and consumers instead tend to rely on a two-stage decision process: first screening a large set of products to identify a promising shortlist, and then comparing those alternatives more carefully before committing to a purchase (Häubl & Trifts, 2000).

Recommender systems serve as interactive decision aids that support consumers at both stages (Häubl & Trifts, 2000). At their core, they are algorithms that estimate the likelihood that a user will find a given item relevant and use that estimate to present the most suitable option from a wide range of possibilities (Dadoun et al., 2021).

Recommender systems are by now a familiar part of everyday digital life. For example, when Netflix suggests a show you end up watching for three hours, or Amazon reminds you that people who bought what you just bought also tend to buy something else, these are recommender systems at work: algorithms estimating what you are likely to want (based on your past search history or peer group consumption) before you have asked for it (Gomez-Uribe & Hunt, 2016; Linden et al., 2003). In each case, the system does the searching so the user does not have to. These are traditional recommender systems. In the airline context, it works the same way, except the stakes are higher: choosing the wrong Netflix show costs you an evening; choosing the wrong flight can cost you considerably more. Originally developed to address information overload in e-commerce and entertainment, recommender systems have since expanded into the travel sector, where the decision environment is complex (Dadoun et al., 2021).

Two kinds of recommender systems need to be told apart here. Traditional recommender systems, like the engines behind Netflix and Amazon, rank items by how relevant they are likely to be, but they hand the user a finished answer: you see what is recommended, not why (Linden et al., 2003; Gomez-Uribe & Hunt, 2016). Recommender systems built on large language models can do something more. In addition to ranking the options, they can explain their choice in plain language, telling the user which criteria mattered and why other options were not selected. It is this kind of recommender system that the present study focuses on. The main topic this study investigated is whether AI explaining its own choice changes how consumers react to it. The underlying recommendation remains the same across all experimental conditions. The only thing that changes is whether the AI explains itself or not (see Chapter 3.2).

In the airline context specifically, recommender systems have the potential to support consumers across the entire traveler journey (see Figure 1) (Dadoun et al., 2021).



(Figure 1, the airline customer journey and the potential impact of recommender systems, Dadoun et al., 2021).

At the shopping stage (and partly the booking stage), which is the focus of this study, the consumer is in a planning mindset and actively weighing trade-offs between available flights (Dadoun et al., 2021). A recommender system at this stage effectively collapses the two-stage decision process described earlier into a single interaction, presenting a recommendation framed as the best match for the user's needs.

Despite this potential, the application of recommender systems in airline retailing remains underdeveloped (Dadoun et al., 2021). Existing research on flight recommender systems has focused almost entirely on the technical side, consisting of improving delay prediction accuracy, modeling user preferences using behavioral data, and optimizing ranking algorithms (Kan et al., 2024). What has received far less attention is how consumers respond to these recommendations once presented to them. Questions about whether users trust the recommendation, how the framing affects their risk perception, and whether they intend to act on it remain largely unanswered (Dadoun et al., 2021). This matters because a technically accurate recommendation is only useful if the consumer is willing to follow it. This study examines two factors that shape that willingness: whether the AI explains the reasoning behind its recommendation (transparency), and the price level of the recommended flight.

Flight booking is a particularly suitable context for studying how AI explanation interacts with price for three main reasons. First, it spans a wide and credible price range within a single product category: the same task, booking a one-week trip, can cost a few hundred or several thousand SEK, depending only on the route, which allows price to be varied without leaving the main category or goal (a one-week vacation) (see Chapter 3.2). Second, the decision is characterized by a high degree of uncertainty for the consumer. Fares, layovers, different airlines with different on-board products are hard to evaluate, and outcomes are non-refundable in the cheapest category, so explanation can reduce uncertainty before commitment. Third, AI-assisted booking is a real and fast-growing commercial development (Zanpure, 2025), which makes the question immediately relevant to airlines and online travel agencies.

The mechanism this study investigates is not specific to air travel. Wherever an AI system recommends a single option from a large, complex set, and the financial stakes vary, the same question arises. Selecting an insurance or mortgage product, choosing among investment options in an online app, or accepting a marketplace listing all share the same structure: a consumer who cannot easily verify the recommendation, a meaningful amount at stake, and an AI that can explain its choice. Flight booking is studied here as a concrete instance of this broader class of high-stakes, AI-assisted consumer decisions.

1.2 Purpose and Research Question

Many studies within the recommender systems literature have demonstrated the positive effect of transparency on trust and purchase intention (Tintarev & Masthoff, 2007; Glikson & Woolley, 2020; Wang & Yin, 2021). Yet other studies in related contexts have also identified conditions under which transparency does not affect, or even harms, such desirable outcomes (Park & Yoon, 2025; Petty & Cacioppo, 1986; Scharowski et al., 2022).

This study aims to extend the literature by introducing price as another such condition. The rationale is inspired by Cox and Rich (1964), who argue that consumers seek information to reduce the perceived risk associated with purchase decisions. In the context of AI recommender systems, transparency can be viewed as an information-provision mechanism that reduces uncertainty. Building on this perspective, the authors argue that perceived risk increases when more is at stake and when uncertainty surrounding a decision is higher. Consequently, as prices increase, transparency should become more valuable in reducing uncertainty, thereby strengthening its effect on purchase intention.

The study also draws on more recent TAM and UTAUT literature to explore the psychological mechanisms underlying the relationship between transparency and purchase intention and to examine whether the moderating effect of price extends to this relationship.

To address these objectives, the following research questions are posed:

RQ1: Does transparency affect purchase intention?

RQ2: Does transparency have a stronger effect on purchase intention when price is high?

RQ3: Through which psychological mechanisms, specifically trust and perceived risk, does transparency influence purchase intention?

These questions are addressed through a controlled 2×2 between-subjects experimental survey design, described in detail in Chapter 3.

1.3 Research Gap

How transparency behaves under high vs. low financial stakes is theoretically important and practically relevant (see chapter 1.2), yet, to the best of the authors' knowledge, it has not been tested in a consumer recommender system purchase context.

Recent work treats transparency as a positive force. Tintarev and Masthoff (2007) identify trust as one of seven core aims of explanations. The positive framing is not universal. Scharowski et al. (2022) document conditions under which explanations fail to raise trust and can even lower it. What remains untested is whether the effect of transparency on trust and behavioural intention differs across price levels. To the best of the authors' knowledge, no study has directly compared transparency under high- and low-stakes conditions in a consumer-facing booking context. This gap motivates the present study.

A second gap concerns the level of transparency studied. Werz et al. (2024) draw a useful distinction between global transparency (openness about the algorithm and the system as a whole) and local transparency (the active display of reasoning behind a single recommendation, often labeled explainability). Most consumer-context studies bundle these into a single transparency construct, which makes it difficult to isolate the effect of explanation-level transparency on consumer responses (Miller, 2019; Scharowski et al., 2022). This study targets local explainability directly (see chapter 1.5), so the effects reported here speak to the specific content of transparency based on explanations.

1.4 Delimitations

The study is delimited along four dimensions.

First, a fictitious AI booking scenario is presented. The result is higher internal validity at the cost of ecological validity, a trade-off described by Shadish et al. (2002) as common in such experimental designs.

Second, the sample is a consumer panel aged 20–39. This age range was set to limit age-related variance in the experiment's main variables, given the sample size constraint.

Third, perceived risk is narrowed to the financial dimension of Jacoby and Kaplan's (1972) framework. Performance, psychological, and social risk dimensions are excluded. Price most directly drives financial risk. This study primarily measured financial risk, complemented by a single item capturing general risk (see chapter 3).

Fourth, transparency is operationalized through explainability. Transparency is implemented through AI-generated explanations of the recommendation logic (see chapter 2.3). Other forms of transparency, such as algorithm-level openness, are not tested. The reasoning is that explainability is the form of transparency most relevant to the lay user (Werz et al., 2024; Miller, 2019). Findings should be interpreted as evidence on explanation-based transparency, not on transparency in a broader sense.

1.5 Structure of the Thesis

This thesis is structured into six chapters. Chapter 2 presents the theoretical framework, reviews the three key models of technology acceptance (TAM, UTAUT, and S-O-R), and develops hypotheses grounded in the literature. Chapter 3 details the methodology, including the 2×2 experimental research design, the independent variables (price level and AI transparency), the survey structure, the pre-study, the sample, and the analytical strategy. Chapter 4 reports the results of the manipulation checks, main effects, and mediation analyses. Chapter 5 discusses these findings in relation to the theoretical framework and the broader literature on AI-assisted decision-making in airline retailing, and offers limitations and directions for future research. Chapter 6 offers concluding remarks.

2. Theoretical Framework

The theoretical framework is built in two parts. Part 1 (section 2.1) reviews the TAM (technology acceptance model) literature to identify the key mechanisms involved in how humans interact with technology in the context of purchasing and introduces the S.O.R model to ground the causal chain between the stimuli, the mechanisms, and the dependent variables. The purpose is to ground the study in proven frameworks and identify what each contributes to the model developed here. Part 2 (sections 2.2–2.7) develops the study-specific theoretical model and the hypotheses.

2.1 Conceptual foundation

2.1.1 Technology Acceptance Model

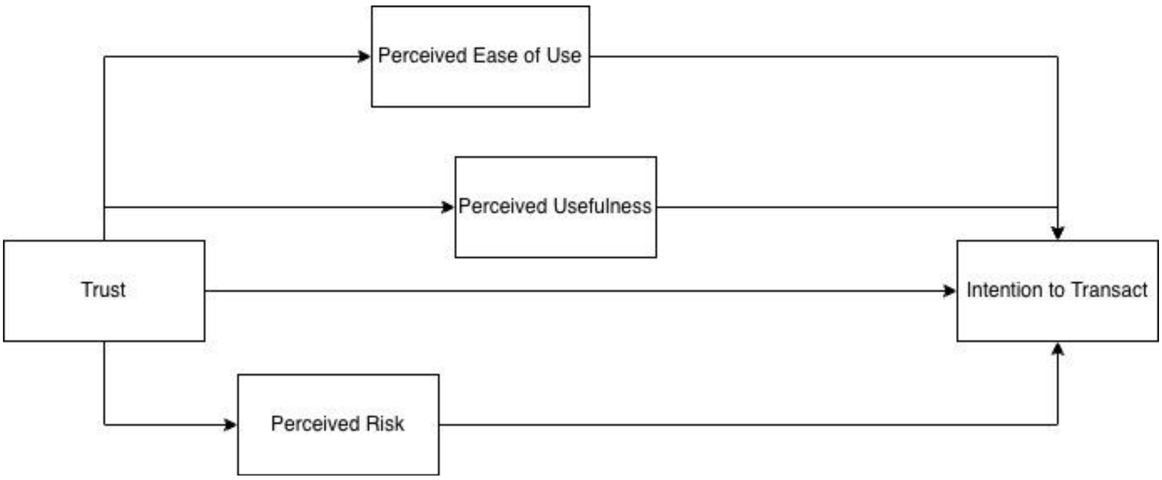
The Technology Acceptance Model (TAM) was introduced by Davis (1989) to explain user acceptance of information technology (see Figure 2). TAM identifies two key beliefs as central drivers of behavioural intention: Perceived Usefulness (PU), the degree to which a person believes a system improves their performance, and Perceived Ease of Use (PEOU), the degree to which using the system is expected to require little effort.

Building on TAM and seven adjacent acceptance models, Venkatesh et al. (2003) developed the Unified Theory of Acceptance and Use of Technology (UTAUT), which organizes acceptance around four determinants: Performance Expectancy, Effort Expectancy, Social Influence, and Facilitating Conditions. UTAUT further highlights the importance of individual differences such as age, gender, and experience in shaping technology acceptance.

In the context of online shopping and e-commerce, both models were also extended to include perceived risk and trust. Gefen et al. (2003) introduced trust as a key variable affecting behavioural intention both directly and through perceived usefulness. Pavlou (2003) built on these findings by introducing perceived risk and its negative relationship with behavioural intention, and by re-establishing trust as a predictor of behavioural intention and perceived usefulness in addition to its effect on the newly introduced risk variable (Gefen et al., 2003; Pavlou, 2003). These two variables were later incorporated into UTAUT as well (Slade et al., 2015; Wanner et al., 2022).

Pavlou's (2003) extension is particularly relevant to the present study, as it focuses on the role of uncertainty in online purchase decisions. Pavlou argues that the distant and impersonal nature of online transactions creates uncertainty around the purchase process, making trust and perceived risk central elements of the decision to purchase. This aligns closely with the logic adopted in the present study, in which transparency is viewed as a mechanism to increase certainty before a purchase decision is made. As such, Pavlou's model serves as a useful foundation for identifying the determinants of purchase intention examined in this study.

The AI recommender context, however, introduces additional considerations beyond those captured by Pavlou's model. The antecedent variables used by Pavlou primarily capture pre-existing beliefs and experiences that consumers bring into an online transaction. In the present study, these factors are represented through initial openness toward AI and prior AI experience, which reflect consumers' pre-existing attitudes toward and familiarity with AI systems (see Chapter 2.7). Age and gender are also included, given their established roles in technology acceptance research. Conversely, perceived ease of use is omitted because the present study focuses on transparency and decision-making rather than system usability. Together, these adaptations provide the foundation for the model developed in this study.



(Figure 2, TAM extension for electronic commerce (Pavlou, 2003)).

2.1.2 Stimulus-Organism-Response Model

The Stimulus-Organism-Response framework (Mehrabian & Russell, 1974; see Figure 4) fits this study for one main reason: the chain it describes matches the chain the study tests. Price and transparency are the stimuli, trust and perceived risk are the internal states they activate, and purchase intention and perceived usefulness are the responses.

(Figure 3, S-O-R framework by Mehrabian & Russell (1974)).

In this study, the price of the recommended flight and the transparency of the AI's reasoning are the stimuli. These activate internal evaluative states: how much the consumer trusts the recommendation and how risky the purchase feels. These internal states determine whether the consumer is willing to book (Mehrabian & Russell, 1974).

S-O-R has been applied to intention outcomes similar to those in this research. Animesh et al. (2011) apply S-O-R to virtual-world commerce in MIS Quarterly, with technological, social presence, and spatial environment features as the stimulus, telepresence and flow as the organism, and intention to purchase virtual products as the response. Kim et al. (2020) extend S-O-R to virtual reality tourism, modeling authentic experience as the stimulus, cognitive and affective responses as the organism, and intention to visit the destination shown in VR as the response. The fit is imperfect: neither study isolates a written explanation as part of the stimulus, and neither manipulates the price of the recommended purchase. The S-O-R chain is still useful because it captures the basic structure of the study design, and because the two cited applications confirm that the chain works in digital purchase contexts where the consumer never touches the product before deciding, with an intention variable as the main outcome.

2.2 Dependent Variables: Purchase Intention and Perceived Usefulness

2.2.1 Purchase Intention

Definition

Purchase intention is the consumer's subjective probability of performing a specific purchase behaviour in the near future (Fishbein & Ajzen, 1975; Ajzen, 1991). In recommendation system research, purchase intention reflects the likelihood that a user will act on a recommendation by completing the transaction it suggests (Pu et al., 2011; Wang & Benbasat, 2009). In the present study, the consumer's stated their likelihood of booking the flight that the AI assistant has recommended.

What drives Purchase Intention

Trust and perceived risk have been established as key determinants of purchase intention in the context of e-commerce. Trust has been identified as both an antecedent of perceived usefulness and a direct predictor of behavioural intention in online purchase contexts characterized by information asymmetry (Gefen et al., 2003; Pavlou, 2003). When trust is high, consumers are more willing to accept the vulnerability of an online transaction and to commit (Pavlou, 2003; Jarvenpaa et al., 2000; Gefen et al., 2003). Perceived risk, on the other hand, has consistently been shown to lower purchase intention. When the expected cost of a wrong decision is high, consumers raise the threshold of confidence required before committing, and non-action becomes the default (Mitchell, 1999; Pavlou, 2003).

Purchase Intention and Purchase Behaviour

A persistent gap exists between behavioural intention and actual behaviour. Warshaw and Davis (1985) distinguish between behavioural intention (what one plans to do) and behavioural expectation (what one believes will happen) and find that the two differ systematically. Jarvenpaa et al. (2000) treat stated intention to buy as a reasonable predictor of subsequent purchase behaviour across comparable online shopping sites. The intention measure used in this study should be interpreted in the same way as a strong proxy for actual booking behaviour in the experimental booking context. Sheeran (2002) empirically confirms this in a meta-analysis of 10 prior meta-analyses, reporting a sample-weighted average correlation of $r = 0.53$ between intention and subsequent behaviour. The correlation rose to 0.82 after correcting for dichotomization and measurement error.

2.2.2 Perceived Usefulness

Definition

Perceived usefulness is the degree to which an individual believes that using the system will help them attain gains in job performance (Davis et al., 1989). Within the technology acceptance literature, perceived usefulness is often seen as an antecedent of behavioural intention, alongside ease of use (see Chapter 2.1).

What drives Perceived Usefulness

In the recommender systems and online purchasing literature, a key variable shown to influence perceived usefulness is trust. Gefen et al. (2003) argue that trust is needed to establish the vendor's credibility in delivering on promises, and, by extension, the consumer's ability to accomplish their task on a website. Pavlou (2003) similarly argues that in the absence of trust, the consumer is less inclined to expect any utility gains from using an interface.

2.3 Transparency

The term "transparency" is itself a container term, meaning different things to different people (Fox, 2007; Miller, 2019). What transparency means in a specific situation often remains poorly defined. This study uses the definition developed by Wanner et al. (2022), who define transparency as the system's ability to explain and reveal its decision rationale to the user through visual means.

2.3.1 Explainability as a Form of Transparency

For a lay user, the internal mechanics and algorithmic complexities behind an explanation are often of secondary importance. Werz et al. (2024) highlight that users, who typically lack a technical background, prioritize whether the information provided is sufficient to clarify why a specific recommendation was made in their case. This shift in focus emphasizes the necessity of "local transparency", which prioritizes the logic behind individual outputs over the broader architectural "global transparency" of the system. This necessity leads to the related concept of explainability, a mechanism that increases transparency (Werz et al., 2024). Explainability is a specific form of transparency, namely the active display of reasons for an AI's decisions (Miller, 2019; Barredo Arrieta et al., 2020).

2.3.2 What constitutes a “good” Explanation?

While the definition of explainability is broad and what constitutes a "good" explanation differs between studies, Miller (2019) identifies four key characteristics of an effective explanation:

Contrastive Nature: People rarely ask why something happened in a vacuum. They ask why it happened instead of something else. A user does not ask why event X happened; they ask why X happened and not Y.

Selective Bias: People do not want the full causal chain. They pick one or two main causes from many possible ones, and the causes they pick depend on their own mental shortcuts and biases.

Causal Over Statistical: AI runs on probabilities, but probabilities rarely satisfy people. A statistical pattern is less useful to a human than a clear cause-and-effect story. The most statistically likely explanation is often not the most useful one.

Social Interaction: Explanations are a social form of knowledge transfer. They function as a conversation in which information is tailored based on the explainer's perception of the listener's existing beliefs.

2.3.3 The Effect of Transparency on Purchase Intention and Perceived Usefulness

Much of the literature on transparency and purchase intention is consistent in finding a positive effect of transparency on purchase intention (in the context of TAM and UTAUT). Wang and Yin (2021) found that AI explanations positively affected decision-making quality and willingness to follow recommendations. Tintarev and Masthoff (2007) similarly showed that providing explanations for recommender outputs increased user acceptance. Sinha and Swearingen (2002) provide evidence that consumers prefer recommendations from systems whose reasoning they can see. Chen et al. (2024) found that the presence of an explanation raised purchase intention by increasing interpretability. Transparency also plays a key role in both increasing trust (see chapter 2.4) and lowering perceived risk (see chapter 2.5), both of which are strong predictors of purchase intention (see chapter 2.2.1). The following is therefore hypothesized:

H1: Higher transparency leads to higher purchase intention.

Similarly, trust is a determinant of perceived usefulness in e-commerce (see Chapter 2.2.2). By increasing trust, transparency would therefore also positively affect perceived usefulness (Wanner et al., 2022).

H2: Higher transparency leads to higher perceived usefulness.

2.4 Trust

2.4.1 Definition and Key Dimensions

Trust is the consumer's willingness to accept vulnerability based on positive expectations about another party's actions (McKnight et al., 1998, 2002). In this study, trust captures the consumer's confidence in the AI booking assistant and in the specific flight recommendation it produces. McKnight et al. (1998, 2002) decompose trust into three trusting beliefs: competence (the AI can provide accurate recommendations), benevolence (the system is oriented toward the user's interests), and integrity (the system is honest and does not mislead the user).

2.4.2 The effect of Transparency on Trust

While evidence on the effect of transparency on trust is mixed, the broader consensus supports the view that transparency has a positive effect on trust. By outlining the logic behind how a recommendation was made, a system reduces the user's uncertainty and increases the perception of the system as competent (Rahman, 2025; Cramer et al., 2008).

The act of disclosing the system's logic serves as a signal of honesty and reliability (Park & Yoon, 2025). Kizilcec (2016) further found that transparency serves as a means to restore trust when user expectations have been violated. He also argued, however, that the effect of transparency on trust is not without limits. While identifying a positive relationship between transparency and trust, he found that a lengthy explanation eroded trust rather than building it. Schmidt et al. (2020) further show that transparency reduces trust when the explanation exceeds users' ability to comprehend, via the mechanism of cognitive overload. Both Scharowski et al. (2022) and Werz et al. (2024) provide further evidence of boundary conditions in which the effect of transparency on trust is reversed, and underscore the importance of the type and form of explanation in shaping this relationship.

In this study, in the absence of prior interaction with the fictional tool used in the experiment, transparency may serve as a signal of both honesty and competence. Furthermore, the recommendation is provided in the context of an online purchase decision, where the transaction's distant, impersonal nature creates uncertainty about the purchase process (Pavlou, 2003). Pavlou (2003) argues that mechanisms that reduce uncertainty can facilitate trust in online transactions. As such, transparency is expected to positively affect trust. Additionally, trust is particularly important as a determinant of purchase intention and perceived usefulness in situations characterized by greater uncertainty (Pavlou, 2003). It can therefore be argued that trust mediates the relationship between transparency and the dependent variables, purchase intention, and perceived usefulness.

H3: Higher transparency leads to higher trust.

H4: The relationship between transparency and purchase intention is mediated by trust.

H5: The relationship between transparency and perceived usefulness is mediated by trust.

2.5 Perceived Risk

2.5.1 Definition and Key Dimensions

Perceived risk is a central concept in consumer behavior and has been widely studied across both traditional marketing and digital contexts for over fifty years. Originally introduced by Raymond A. Bauer (1960, as cited in Mitchell, 1999), the concept refers to the uncertainty and potential negative consequences associated with a purchase decision. In modern literature, perceived risk is generally understood as a multidimensional concept comprising financial, performance, physical, psychological, and social risks (Jacoby & Kaplan, 1972).

Financial risk refers to the potential for monetary loss associated with a purchase decision (Jacoby & Kaplan, 1972). As this study specifically examines the effect of transparency under varying price conditions, financial risk is deemed the most relevant dimension of perceived risk, given its direct conceptual link to price, and will therefore be emphasized throughout the paper.

2.5.2 The effect of Transparency on Perceived Risk

Cox and Rich (1964) emphasize perceived risk as a function of two factors: the amount at stake and the level of subjective certainty regarding the outcome. They argue that consumers typically seek to reduce perceived risk before making a purchase by increasing certainty. The authors further identify two main types of consumers based on the form of risk reduction they adopt: those who rely on prior experience and those who rely on external information. In the context of AI recommender systems, transparency, through explainability, serves the latter purpose. Dabholkar and Sheng (2012) make a similar argument but for consumer participation rather than transparency in the context of recommender systems. They also establish the negative effect of financial risk, specifically on purchase intention. These arguments, combined with the established relationship between perceived risk and purchase intention in general (see Chapter 2.2.1), indicate that perceived risk mediates the relationship between transparency and purchase intention.

H6: The relationship between transparency and purchase intention is mediated by perceived risk

2.6 Price and Transparency

Price plays a central role in defining how a consumer approaches a purchasing decision.

First, price is closely tied to perceived financial risk. It is one of the two variables in the perceived risk equation defined by Cox and Rich (1964), with perceived risk being a central determinant of purchase intention (Dabholkar & Sheng, 2012). It follows that a price increase increases the need for a risk-reduction measure to lower uncertainty before the consumer engages in purchasing behavior. Transparency is therefore expected to have a stronger effect on purchase intention and perceived usefulness in the high-price context.

H7: Higher transparency leads to higher purchase intention when price is high.

H8: Higher transparency leads to higher perceived usefulness when price is high.

Second, Dabholkar and Sheng (2012) argue that when price (financial risk) is low, a higher level of consumer participation is likely to be viewed as unnecessary effort and is unlikely to increase trust. A similar logic applies to transparency in this study. Chapter 2.4.2 highlights how transparency can negatively affect trust through cognitive overload and specifically cites Kizilcec (2016), who found that lengthy explanations eroded trust rather than building it. Combining these two arguments, while transparency would lead to higher levels of trust when needed in a high-price context, it would also lead to lower levels of trust when lengthy and unnecessary in a low-price context. Price is therefore expected to moderate the relationship between transparency and trust.

H9: The relationship between transparency and trust is moderated by price.

Third, since transparency is expected to be most relevant when the financial stakes of the decision are high, the proposed psychological mechanisms are likewise expected to be most relevant in high-stakes contexts. Trust and perceived risk are therefore conceptualized as mechanisms through which transparency may influence consumer responses, particularly when consumers face substantial financial risk. In low-price contexts, the need for additional certainty is weaker, meaning that transparency is less likely to meaningfully affect trust or perceived risk. Consequently, the indirect effects proposed in H4–H6 are expected to be most relevant under high-price conditions.

2.7 Control variables

To isolate the effects of transparency and the other focal constructs in the model, it is necessary to control for individual differences known to influence technology adoption and online purchasing behavior. In the context of recommender systems, artificial intelligence, and technology adoption more broadly, such individual differences play a key role in shaping how users interact with specific technologies. Pavlou's (2003) model of online purchasing includes three individual-difference-related control variables: web retailer reputation, satisfaction with prior online transactions, and web-shopping frequency. These variables capture important aspects of experience and trust in e-commerce settings. However, because the focal object in this study is an AI recommender system rather than a specific retailer or online channel, these controls are not sufficient on their own.

AI is a polarising technology: many users express concerns that are directed at the technology itself rather than at particular firms or platforms. Some consumers are highly open to adopting AI, while others are reluctant, and this initial openness toward AI substantially shapes how they interact with AI-based systems. A more positive baseline level of openness can lead users to follow AI recommendations even when explanations are minimal. In contrast, lower openness can hinder meaningful interaction with the AI system, fostering skepticism and lower trust. Recent work on trust and attitudes toward AI shows that global AI-related predispositions influence how individuals evaluate and engage with AI systems (Afroogh et al., 2024). In line with this perspective, the present study includes initial openness toward AI as an AI-specific control variable. Conceptually, it captures a baseline stance toward AI-powered technologies that complements, rather than replaces, the more traditional e-commerce controls proposed by Pavlou (2003).

Beyond AI- and e-commerce-specific controls, it is also important to account for general differences in technology adoption identified in models such as TAM and UTAUT. Over time, these frameworks have highlighted four key individual differences that shape relationships between constructs and influence users' attitudes and behaviours: age, gender, prior experience, and voluntariness of use (Venkatesh et al., 2012). In this study, age and gender are retained as control variables, and prior experience with AI recommender systems for online purchases is also included to capture the role of usage frequency, analogous to Pavlou's (2003) web-shopping frequency. Voluntariness of use is omitted, as the use of AI recommender systems for flight purchases in this context is inherently voluntary and thus does not vary across respondents.

Taken together, the analyses control for age, gender, prior experience with AI recommender systems for online purchases, and initial openness toward AI, ensuring that these individual differences do not confound the estimated effects of transparency and the other focal constructs.

2.8 Summary of the model

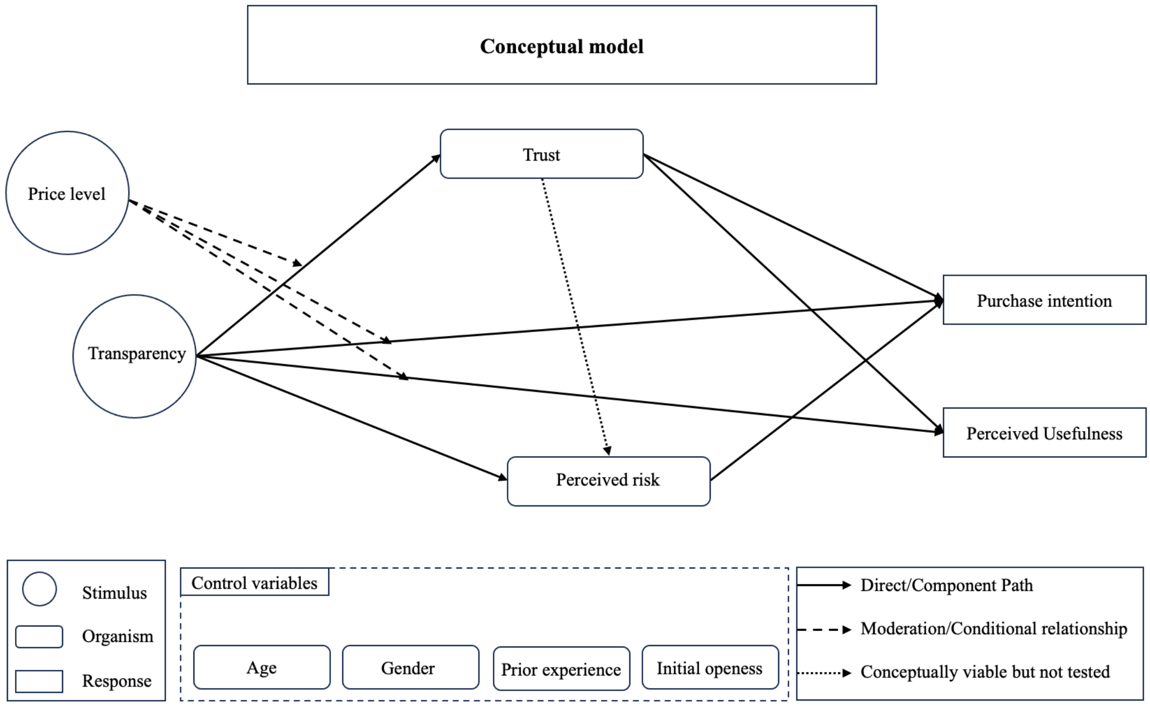
This study tests how the effect of transparency on purchase intention and perceived usefulness varies with price level.

The logic of the study design draws mainly on the S-O-R model, which provides the basic causal chain: transparency and price as the stimuli, trust and perceived risk as the organisms, and purchase intention and perceived usefulness as responses. TAM literature and Pavlou's (2003) extension of TAM to e-commerce were then used to identify key psychological mechanisms driving the effect of transparency and to examine how the study's variables interact with one another. UTAUT literature was used to identify relevant control variables. The literature on the relationship between perceived risk and purchasing behavior was then used to hypothesize differences in the effects of transparency across price levels. Specifically, transparency and the psychological mechanisms through which it operates are expected to be most relevant in high-price contexts, where consumers face greater financial risk and uncertainty. This leads to the following hypotheses:

Hypothesis	**Statement**	**Type**
H1	Higher transparency leads to higher purchase intention.	Direct effect
H2	Higher transparency leads to higher perceived usefulness.	Direct effect
H3	Higher transparency leads to higher trust.	Direct effect
H4	The relationship between transparency and purchase intention is mediated by trust.	Mediation
H5	The relationship between transparency and perceived usefulness is mediated by trust.	Mediation
H6	The relationship between transparency and purchase intention is mediated by perceived risk.	Mediation
H7	Higher transparency leads to higher purchase intention when price is high.	Conditional effect
H8	Higher transparency leads to higher perceived usefulness when price is high.	Conditional effect
H9	The relationship between transparency and trust is moderated by price.	Moderation

Table 1, Hypotheses.

If transparency only affects purchase intention and perceived usefulness (the dependent variables) in the high-price condition, then the mediation hypotheses (H4, H5, and H6) are only expected to hold in that condition and will therefore be tested exclusively in the high-price condition.



(Figure 4, conceptual model)

3. Methodology

3.1 Research Design

This study adopts a quantitative, experimental research design, following the deductive approach described by Bryman and Bell (2011), in which hypotheses are derived from existing theory and then tested through structured empirical data collection (see Chapter 2 for hypothesis development). A deductive approach is appropriate here because the study builds directly on established theoretical frameworks, specifically the Technology Acceptance Model (TAM), trust theory, and perceived risk literature, and seeks to test how specific manipulations influence a set of dependent variables (see chapter 2). A quantitative strategy is therefore the natural choice: the goal is to measure correlations among variables in a sufficiently large sample and draw conclusions about their relationships (Bryman & Bell, 2011; Broström & Wennberg, 2025). A qualitative approach was deliberately not used, as qualitative methods are not designed to test causal hypotheses across groups or to generalize findings to a broader population, but rather to develop theory through an inductive approach (Bryman & Bell, 2011).

More specifically, the study employs a 2×2 between-subjects experimental design, in which each participant is randomly assigned to one of four conditions and exposed to only that condition throughout. The design manipulates two independent variables, price (low vs. high) and transparency level (low vs. high), resulting in four distinct scenarios. This structure allows both the main effects of each independent variable and their potential interaction to be examined within a single study (Söderlund, 2018). The Central Limit Theorem further supports the use of parametric tests. As long as the sample per cell is reasonably large, the sampling distribution of means will approximate normality even for the ordinal Likert data used throughout this survey (Broström & Wennberg, 2025). The research context, AI-assisted flight booking, positions the study within the literature on algorithmic recommender systems (see chapter 1.1) and technology acceptance (see chapter 2.1) (Gefen et al., 2003). Experiments are particularly well suited to this context because random assignment eliminates pre-existing differences between groups, allowing observed differences in the dependent variables to be attributed to the manipulations themselves (Shadish et al., 2002; Söderlund, 2018).

3.2 Experimental Design

Participants were randomly assigned to one of four conditions arising from the 2×2 crossing of price level and transparency level. In the low-price conditions, the scenario described a flight search from Stockholm to Warsaw, a relatively affordable short-haul route. Under high-price conditions, the destination was Rio de Janeiro, which represented a more expensive long-haul flight (see Appendix A). The consumer's subjective evaluation of a price as high or low, referred to here as perceived price level, is distinct from the objective price point and is verified empirically through the manipulation check item discussed in chapter 3.3 (Völckner, 2008).

3.2.1 Price manipulation and validity considerations

This operationalization has implications for construct validity that must be acknowledged. The destination swap means that price level is not manipulated in isolation: Warsaw and Rio de Janeiro differ not only in price, but also in travel distance, flight duration, connection complexity, and potentially in the degree of unfamiliarity a Swedish consumer may have with the destination. Thus, price is not manipulated in complete isolation from other route characteristics.

An alternative design was considered and ultimately rejected for two reasons. Presenting two unreasonably different prices for an identical product risks undermining the credibility of the scenario: participants familiar with online flight booking are unlikely to find it plausible that the same route, airline, and fare class could cost, for example, 800 SEK and 6,000 SEK simultaneously. This would undermine ecological validity. Furthermore, the price variation observed in real flight booking environments is typically tied to route characteristics such as distance, airline competition, and seasonality, rather than to arbitrary price-point changes for an otherwise identical flight. A design that reflects this natural covariance is therefore more representative of how consumers actually encounter price in this context, even at the cost of perfect construct isolation (Shadish et al., 2002; Cook & Campbell, 1979). Consequently, the effects attributed to price in this study should be interpreted with caution, as they may partly reflect correlated factors such as complexity, destination familiarity, and travel distance. At the same time, in flight booking and many other markets, price is naturally confounded with such factors, so the design mirrors the way consumers encounter prices in practice.

3.2.2. Transparency manipulation and stimuli

Transparency was manipulated by varying the information the AI assistant communicated alongside its recommendation. In the low-transparency conditions, the flight result was displayed without any explanation of the AI's reasoning and was accompanied by a short filler text describing the destination. In the high-transparency conditions, the AI also provided a brief explanation of how it had weighted price and travel time. Each condition began with an identical introductory text, ensuring that any differences in responses could be attributed to the experimental manipulations rather than to differences in framing. When groups are equivalent at baseline, post-manipulation differences in the dependent variables can be attributed to the manipulations themselves (Shadish et al., 2002; Cook & Campbell, 1979).

The booking-portal screenshots used as visual stimuli followed common practice in experimental research on digital interfaces (e.g., Werz et al., 2024; Rosenberger et al., 2025). All screenshots were created in Lovable using an identical visual layout across conditions (see Appendix B for the prompt). The only differences between conditions were the destination, the price, and the presence or absence of the AI explanation.

3.2.3 Pre-studies and refinement of transparency manipulation

A first pre-study (N = 30) indicated that the initial transparency manipulation did not produce a significant difference in perceived transparency, prompting a revision of the high-transparency text. To strengthen the manipulation, a new structured prompt was developed in Gemini, based on transparency criteria derived from Miller (2019) and Werz et al. (2024) and outlined in section 2.4.2. The prompt required the explanation to contrast the recommended flight against alternatives, list only selection criteria relevant to the user's stated preferences, and cite the data sources consulted (see Appendix C). Using this prompt, a revised high-transparency explanation was first generated for the high-price flight. The AI was then instructed to produce the explanation for the low-price flight by using the high-price explanation as a template, so that the low-price explanation was structurally identical to the high-price one in terms of selection criteria, level of detail, and comparative structure, with only content adjustments needed to reflect the different route.

After the revised explanations were generated, the filler text used in the low-transparency conditions was adjusted to match the explanations in length, ensuring that any observed differences between conditions would reflect the presence or absence of AI reasoning rather than differences in text length (see Appendix D). All texts were checked for realism and factual correctness before inclusion.

A second pre-study ($N = 23$) confirmed that the revised stimuli produced a large and statistically significant difference in perceived transparency overall (diff = 1.71, $t(21) = 3.44$, $p = .003$, $d = 1.44$, 95% CI [0.50, 2.35]) and in the high-price condition (diff = 2.00, $t(9) = 2.56$, $p = .031$, $d = 1.55$, 95% CI [0.14, 2.90]). In the low-price condition, the effect size was similarly large ($d = 1.30$, 95% CI [−0.01, 2.55]), although the difference narrowly missed conventional significance (diff = 1.53, $t(10) = 2.21$, $p = .051$), likely due to the small subgroup size. Overall, these results indicated that the revised transparency manipulation functioned as intended before the main data collection.

3.2.1 Manipulation Checks

To verify that the experimental manipulations operated as intended, the final block of the survey included manipulation-check items. Participants rated how they perceived the flight price (high vs. low) and how transparent they perceived the AI system to be. These measures are used to confirm that the price and transparency conditions induced the expected differences in perceived price level and perceived transparency before drawing conclusions about the effects on the main outcome variables, in line with standard practice in experimental marketing research (Söderlund, 2018).

3.3 Measures

All constructs, except for the screener, the two demographic items, and the attention check, are measured using multi-item Likert scales anchored from 1 (strongly disagree) to 7 (strongly agree), with an additional 'Don't know' option. Multi-item scales are used because they improve reliability (see chapter 3.7) by averaging out random measurement error, and the 'Don't know' option allows respondents to signal genuine uncertainty rather than guessing, which reduces noise in the data (Churchill, 1979; Nunnally, 1978). Likert scales are technically ordinal: the response options have a ranked order, but the psychological distance between adjacent points is not guaranteed to be equal. Treating them as interval-level data, which enables the calculation of means and the use of parametric tests such as ANOVA, is standard and broadly accepted practice in marketing and social science survey research. This assumption is followed throughout this study but is acknowledged as a well-known weakness of the Likert scale (Broström & Wennberg, 2025). Some measures, including perceived risk and trust, were compressed due to survey length constraints.

Perceived Risk (3 items)

Perceived risk is measured using two items capturing financial risk and a third item capturing general risk factors in the context of the AI-assisted booking recommendation. The two financial risk items are adapted from Jacoby and Kaplan's (1972) foundational perceived risk scale.

Trust (6 items)

Trust is measured across six items reflecting the three core dimensions of competence, benevolence, and integrity, consistent with the trust framework established by McKnight (2011) and adapted to the AI recommender context (see chapter 2.4).

Perceived Usefulness (3 items)

Perceived usefulness is measured with three items adapted from Davis (1989) and Venkatesh et al. (2003).

Initial Openness (1 item) and Prior Experience (1 item)

Initial openness to AI and prior experience were both captured using single items as controls for general individual disposition toward AI tools and purchase-specific AI tool usage, respectively.

Purchase Intention (1 item)

Purchase intention was measured through a single item asking the user about the likelihood of purchasing the ticket recommended by the AI assistant. Bergkvist and Rossiter (2007) show that single-item measures are appropriate for constructs with a concrete singular object and a concrete attribute, which applies to purchase intention for a specific AI-recommended flight. Söderlund et al. (2001) further note that single-item measures of intention have dominated the purchase intention literature.

Adoption Intention (4 items)

Items for adoption intention were adapted from prior literature in the contexts of artificial intelligence adoption (Shi et al., 2021) and tourism (Kaushik et al., 2015), in addition to Venkatesh et al. (2012). The construct was collected as a covariate of purchase intention to verify that purchase intention was not sensitive to specific framing. The high correlation between the two measures ($r = .70$, $p < .01$; see Table 7) confirms this view.

Manipulation Checks (4 items)

Four items serve as manipulation checks. Perceived transparency is assessed with three items adapted from Sinha and Swearingen (2002) and Rosenberger et al. (2025), capturing the participant's perceived transparency. One item was dropped from the study to improve the reliability of the scale (Table 3). Perceived price level is assessed with one item adapted from Lichtenstein et al. (1993). To align this study with the usual design of Norstat surveys, these questions were asked toward the end of the survey.

3.4 Survey Structure

The survey opens with a screener question asking whether the participant has ever booked a flight online. Respondents who answer negatively are immediately excluded, as prior booking experience is a prerequisite for the scenario to be meaningful. Following the screener, two brief demographic items collect age and gender. The target age group of 20 to 39 is enforced at this stage, and respondents outside this range are filtered out. This restriction ensures the sample is drawn from the cohort with the highest rates of online travel booking and AI tool adoption, making it directly relevant to the study's context (Eurostat, 2026), while also reducing heterogeneity in prior digital experience that could confound the experimental comparisons (Bryman & Bell, 2011; Shadish et al., 2002; Söderlund, 2018).

After the demographic items, participants are randomly assigned to one of the four experimental conditions via Qualtrics Survey Flow and encounter their specific scenario and screenshot. The question blocks that follow are presented in a fixed order: perceived risk, trust, perceived usefulness, attention check, perceived credibility, individual disposition, and finally the behavioral outcome items, followed by the control variables. This structure aligns with the default design of a Norstat survey. Constructs most directly tied to the scenario are measured first, while general disposition items, behavioral outcomes, and control variables come later, to avoid priming effects that could inflate correlations between constructs (Broström & Wennberg, 2025; Churchill, 1979). Page breaks between construct blocks reduce satisficing and ensure each construct is evaluated as a fresh, separate judgment (Churchill, 1979). An attention check item was embedded midway, and participants who completed the survey in less than 50% of the median completion time were flagged for exclusion, as this duration was considered insufficient for meaningful engagement with the stimulus material (Söderlund, 2018). The straightlining cut-off point was set at 80%, in line with best practices (Kim et al., 2019).

3.5 Sampling and Data Collection

The target population consists of Swedish adults between 20 and 39 years of age who have booked a flight online at least once. This demographic is the most directly relevant group for studying responses to AI-assisted flight recommendations, as it combines high rates of online travel booking with relatively high familiarity with AI tools (Eurostat, 2026; Statista, 2026). This increases internal validity but limits generalizability, a well-known limitation of this study (Söderlund, 2018). Requiring prior booking experience is important because, without it, participants would be unfamiliar with the task context and their responses would reflect confusion rather than genuine psychological reactions to the manipulations (Söderlund, 2018). Data were collected via Norstat, a professional Swedish market research panel, which is a well-established method for reaching a broad demographic range within a target population in experimental consumer research (Söderlund, 2018).

Because a between-subjects design means each respondent sees only one condition, a larger overall sample is required than would be needed for a within-subjects design (Shadish et al., 2002). For a 2×2 factorial design, the literature generally recommends a minimum of 30 to 50 respondents per cell for medium-sized effects (Cohen, 1992). To ensure higher data quality and more stable estimates, this study targeted approximately 80 to 90 participants per condition, giving a total target of around 320 respondents. The survey was administered in Swedish, as this reduces cognitive load and minimizes the risk of scale item misinterpretation, improving construct validity (Churchill, 1979). The survey was originally drafted in English, then translated into Swedish and validated by Swedish-speaking peers and by Norstat to ensure accurate translation.

The main study was administered via the Norstat consumer panel between March 30 and April 14, 2026. A total of N = 669 responses were collected. The data were first cleaned manually using Microsoft Excel. 17 observations were removed based on the age criterion (Age below 20 and above 39), 94 were omitted based on the screening criterion (no prior online flight booking experience), and 25 were excluded for incomplete responses, leaving 533 responses. The second step of cleaning was performed in Stata, removing 44 respondents who failed the attention check, 33 who completed the survey below the minimum time threshold (set at 50% of the median completion time of 220 seconds, following Greszki et al., 2015), and 15 who had a straightlining percentage above 80% (Threshold recommended by Qualtrics (2024)). For the remaining 441 observations, A third cleaning measure was undertaken to remove the 77 participants whose answers to AI1 and AI2 were similar, since these two items are reverse-coded. A final sample of N = 364 remained for analysis.

Randomization check

To verify that randomization was successful across the four groups, a chi-square test of independence was conducted on the distribution of participants across the four conditions. The results indicate that the group sizes did not differ significantly ($\chi^2(1) = 0.35$, $p = .553$). Randomization is therefore deemed successful.

	Low Transparency	**High Transparency**	**Total**
Low Price	103	92	195
High Price	84	85	169
Total	**187**	**177**	**364**

Pearson chi2(1) = 0.352 Pr = 0.553

Table 2, sample analysis.

3.6 Analytical Strategy

The analytical strategy follows two stages. First, the hypotheses are tested on the full sample. Second, if H7 holds and H1a does not, the analysis focuses exclusively on the high-price subsample to examine the mediation paths in the condition where motivation is expected to be strongest.

The analysis then proceeded sequentially. First, descriptive statistics, including means, standard deviations, and item-level distributions, were computed for all constructs to identify any distributional issues before inferential testing. Second, manipulation checks were examined using independent-samples t-tests, comparing the price perception item across price conditions and the transparency perception items across transparency conditions, to confirm that the manipulations worked as designed (Broström & Wennberg, 2025). Third, construct reliability was assessed for each multi-item scale using Cronbach's alpha, with $\alpha \geq 0.70$ as the minimum acceptable threshold (Nunnally, 1978; Churchill, 1979).

The main hypotheses (see chapter 2) were tested using two-way ANOVA, which is the appropriate method for factorial between-subjects designs with two categorical independent variables (Broström & Wennberg, 2025; Shadish et al., 2002). Both main effects and interaction effects were examined. A contrast analysis was then conducted to test H7 specifically and to examine whether transparency had an effect on purchase intention and perceived usefulness. The effect of transparency on trust was assessed in a similar fashion to test hypothesis H2. Mediation and moderation analyses were conducted using the PROCESS macro with bootstrap confidence intervals to handle the non-normal distribution of indirect effects (Preacher & Hayes, 2008).

All analyses were conducted in Stata. R was used specifically for the PROCESS macro and to generate the descriptive statistics and the correlation matrix.

3.7 Validity and Reliability

According to Bryman and Bell (2011), reliability refers to the consistency and repeatability of a measure, while validity concerns whether a measure captures what it is intended to capture. Bryman and Bell (2011) identify several distinct forms of validity relevant to quantitative experimental research: internal validity, external validity, ecological validity, and construct validity. Each is addressed in turn below.

3.7.1 Internal Validity

Internal validity refers to whether observed differences between groups can genuinely be attributed to the manipulated independent variables rather than to other factors (Bryman & Bell, 2011; Shadish et al., 2002). The primary protection in this study is the random assignment of participants to conditions, which distributes pre-existing individual differences equally across groups so that post-manipulation differences are attributable to the conditions themselves (Shadish et al., 2002). A second protection is the use of a fully standardized scenario: the scenario text, booking interface layout, and all survey items are identical across conditions, with only the destination, the price, and the presence or absence of the AI explanation varying. The filler text in the low-transparency conditions matches the length of the explanation text in the high-transparency conditions, controlling for any outcome differences arising from text length (Cook & Campbell, 1979). Details of both texts are provided in the Appendix (see Appendix D). Finally, the attention check and speeding exclusions remove data from participants who were not genuinely engaging, reducing noise in the dataset (Söderlund, 2018).

3.7.2 External Validity

External validity concerns the degree to which findings can be generalized beyond the sample and setting of the study (Bryman & Bell, 2011; Shadish et al., 2002). The use of a professional consumer panel such as Norstat, rather than a convenience sample such as a student population, substantially improves external validity by producing a sample that is more representative of the target population of Swedish online travelers aged 20–39 (Söderlund, 2018). That said, the study's generalizability is bounded by the specific demographic and the hypothetical scenario format. As outlined in the delimitations, findings may not extend directly to older age groups, other cultural contexts, or real booking environments where actual money is spent.

3.7.3 Ecological Validity

Ecological validity refers to whether the study conditions and materials reflect the natural environment in which the behavior would occur (Bryman & Bell, 2011). The scenario format used here, in which participants view a screenshot of a booking interface and respond to hypothetical recommendations, is not a real transaction. The booking portal interface was, however, designed to resemble a professional commercial system, using the Skyscanner design as a reference. The scenario remains hypothetical, and findings should therefore be interpreted with care.

3.7.4 Construct Validity

Construct validity, which concerns whether measures actually capture the theoretical concepts they are intended to represent, is supported in this study by the use of established, previously validated scales from the trust, TAM, and perceived risk literatures (Bryman & Bell, 2011; Churchill, 1979). Each scale has a documented history of use in similar theoretical contexts (see chapter 3.3). Where items were reduced for survey length, the necessary adaptations preserved the original construct definitions. The inclusion of a 'Don't know' response option further supports construct validity by reducing measurement error from forced guessing (Broström & Wennberg, 2025).

One construct validity concern specific to this study is the price manipulation. Because price level is manipulated via destination choice, differences observed between price conditions may partly reflect variation in other components, such as perceived trip complexity or greater unfamiliarity with long-haul flights, rather than financial involvement per se. This limitation is acknowledged in chapter 3.2.

The correlation matrix (see Table 7) supports discriminant validity. Among the focal constructs, correlations ranged from -.56 (between perceived risk and trust in the agent) to .74 (between trust in the agent and perceived usefulness). While some constructs were moderately to strongly associated, these relationships were consistent with the theoretical model and remained below levels that would suggest conceptual overlap. Most importantly, no inter-construct correlation exceeded the recommended threshold of .85 (Kline, 2005), providing further evidence that the constructs are empirically distinguishable. Overall, the correlation pattern supports the distinctiveness of the constructs in this study.

3.7.5 Reliability

Reliability refers to the consistency of a measure: whether it produces stable, repeatable results across different items or occasions (Bryman & Bell, 2011). The most common threat to reliability in survey research is that individual items measuring the same underlying construct may capture slightly different constructs or be interpreted differently by respondents. The standard way to manage this is through multi-item scales, which average out the error associated with any single item and produce a more stable overall representation of the latent construct (Churchill, 1979; Nunnally, 1978). This logic is reflected in the measurement approach throughout the study: all key constructs are measured with multiple items wherever possible. Trust, being a particularly important and multidimensional construct, is measured using six items that span different elements of trust (see chapter 3.3). Perceived risk and perceived usefulness are each measured with three items. Bergkvist and Rossiter (2007) show that single-item measures are appropriate for constructs with a concrete singular object and a concrete attribute, which applies to purchase intention for a specific AI-recommended flight; this is in line with Söderlund et al. (2001). The screener, demographic questions, attention check, and single-item control variables also use a single response.

At the analytical level, reliability was formally assessed using Cronbach's alpha for each scale. Cronbach's alpha measures the degree of inter-item consistency, that is, how strongly the items within a scale correlate with one another. Following the standard threshold established in the psychometrics literature, values of $\alpha \geq 0.70$ are considered the minimum acceptable level for social science research, while values above 0.80 indicate good reliability (Nunnally, 1978; Churchill, 1979). If any scale fell below this threshold, item-level diagnostics were used to identify poorly loading items, which were then considered for removal.

A further consideration is the stability of measurement across the experimental conditions. Because participants are randomly assigned to conditions, and because all survey items are nearly identical across groups, any systematic differences in scale scores between conditions should reflect genuine differences in responses to the situation rather than differences in how the items were understood or administered. This contributes to what Bryman and Bell (2011) refer to as the stability dimension of reliability: the idea that the measurement instrument performs consistently across different groups or contexts. The between-subjects design also means that each respondent encounters only one condition, so there is no risk of reliability being undermined by fatigue or carry-over effects from repeated exposure to the same construct across conditions (Shadish et al., 2002).

Before testing the main hypotheses, the internal consistency of each multi-item scale was assessed using Cronbach's alpha (see chapter 3.6), following the threshold of $\alpha \geq 0.70$ recommended for social science research (Nunnally, 1978; Churchill, 1979). As shown in Table 3, all scales exceeded this threshold, and the majority exceeded 0.80, indicating good to excellent reliability. The trust scale showed the highest internal consistency ($\alpha = .911$). Cronbach's alpha is not the optimal reliability measure for two-item scales (Eisinga et al., 2013). Stata accounts for this by reporting the inter-item correlation, which Eisinga et al. (2013) identify as the appropriate alternative.

Variable	Abbreviation	Number of items	Cronbach's Alpha	Item	Item Alpha
Perceived Transparency[1]	PT	3	0.8827	PT1	0.7829
				PT2	0.7599
				PT3	0.9310
Perceived risk	PR	3	0.8347	PR1	0.7705
				PR2	0.7135
				PR3	0.8275
Trust in the agent	TIA	6	0.9106	TIA1	0.8983
				TIA2	0.8920
				TIA3	0.8935
				TIA4	0.8941
				TIA5	0.8901
				TIA6	0.8990
Perceived usefulness	PU	3	0.8789	PU1	0.8470
				PU2	0.7813
				PU3	0.8533
Adoption intention	AI	4	0.9427	AI1	0.9144
				AI2	0.9313
				AI3	0.9191
				AI4	0.9356

Table 3, scale reliability.

[1] For perceived transparency, PT3 was dropped from the analysis.

3.8 Descriptive statistics

3.8.1 Sample characteristics

3.8.1.1 Demographics

The sample was well balanced in gender, with 51.5% of participants identifying as female and 48.35% identifying as male (n = 2 identified as non-binary). The age distribution was well spread across the target range (20–39), with the exception of the 20–24 age group, which was underrepresented (12.91%) compared to the rest of the sample (see Table 4).

Characteristic	Category	Frequency (n)	Percentage (%)	Cumulative (%)
Age	20-24	47	12.91	12.91
	25-29	109	29.95	42.86
	30-34	99	27.2	70.06
	35-39	109	29.95	100
Gender	Icke-binär	2	0.55	0.55
	Kvinna	186	51.1	51.65
	Man	176	48.35	100

Table 4: demographic characteristics of the sample.

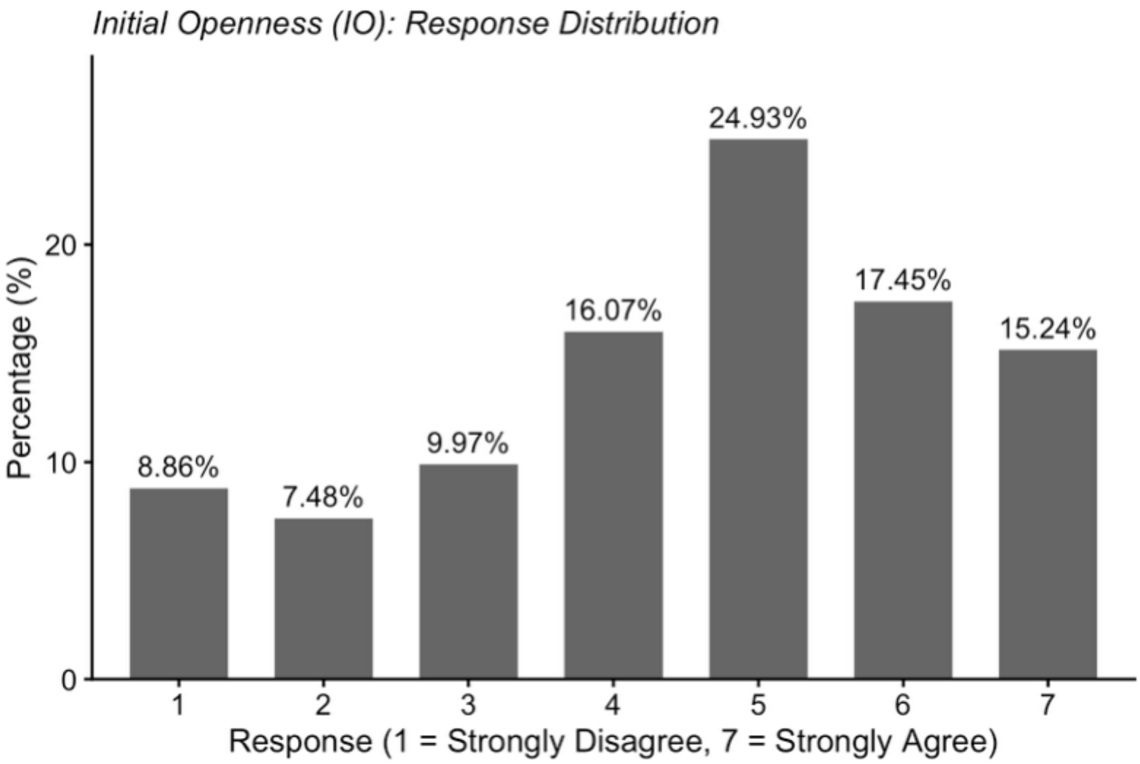
3.8.1.2 Initial Openness and Prior Experience

On average, participants scored higher on IO (M = 4.54, SD = 1.80) than on PRE (M = 2.75, SD = 1.58), suggesting that the sample was receptive toward using AI recommendations, even though prior experience did not match that openness (see Table 5). A more detailed view of each response distribution is shown in Figure 5 for IO and Figure 6 for PRE.

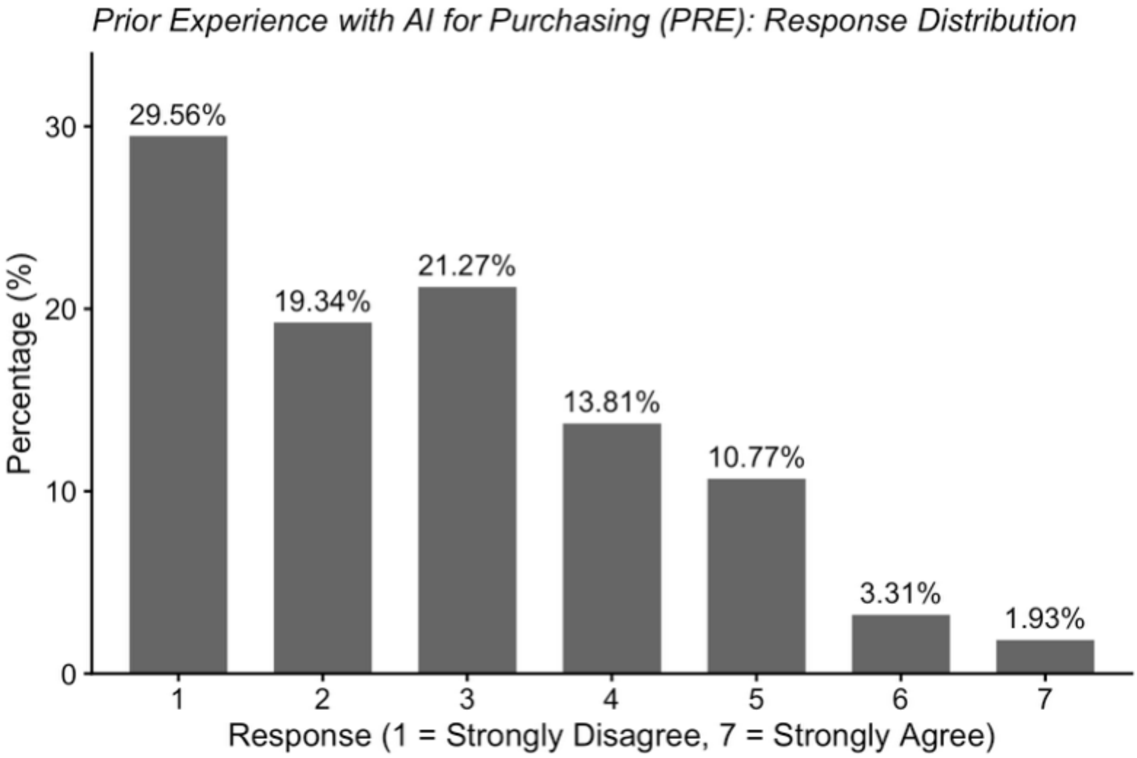
Variable	n	Mean	SD	Min	Max
IO	361	4.54	1.798	1	7
PRE	362	2.746	1.579	1	7

Note. IO = Initial Openness; PRE = Prior Experience with AI. Scale range: 1-7.

Table 5, descriptive statistics for AI experience variables.



(Figure 5, initial openness (IO): response distribution).



(Figure 6, prior experience with AI for purchasing (PRE): response distribution).

Because initial openness toward AI and prior experience with AI recommender systems were measured after exposure to the experimental scenario, they could in principle be influenced by the manipulations, even though random assignment was handled by a professional survey provider (Norstat), which reduces the risk of systematic allocation bias. Both variables were captured through self-reported survey items intended to reflect relatively stable user characteristics. While the underlying dispositions are unlikely to change as a result of a single brief scenario, the reported scores may nevertheless be shaped by how participants experience the manipulation and how they reinterpret their own attitudes and experience in that context. To examine this possibility, the means of both variables were compared across the four experimental groups, and two-way ANOVAs with transparency and price as factors were estimated. For initial openness, the transparency × price interaction was significant, $F(1, 357) = 4.68$, $p = .031$, indicating that openness scores differed across conditions. For prior experience, both the main effect of transparency and the transparency × price interaction were significant, $F(1, 358) = 5.72$, $p = .017$, and $F(1, 358) = 4.95$, $p = .027$, respectively, again showing systematic differences across cells. These patterns suggest that participants' self-reported initial openness and prior experience are likely to have been influenced, at least to some extent, by the transparency × price manipulation, even if their underlying traits remained stable.

Despite this limitation, both variables are included as controls in the main analyses to account for individual differences in AI attitudes and prior exposure to AI recommender systems. However, because they are self-reported measures that are likely to be partially post-treatment, controlling for them may attenuate the estimated main and interaction effects of transparency. The results should therefore be interpreted with caution, as the observed transparency effects may represent conservative estimates of the true impact, particularly in the high-price condition where differences in initial openness and prior experience are most pronounced.

Cell	**n**	**M (IO)**	**SD (IO)**	**n**	**M (PRE)**	**SD (PRE)**
T Low / P Low	102	4.559	1.755	103	2.709	1.512
T Low / P High	83	4.253	1.766	84	2.393	1.456
T High / P Low	91	4.418	1.995	91	2.736	1.632
T High / P High	85	4.929	1.609	84	3.155	1.654

Note. IO = Initial Openness; PRE = Prior Experience with AI for Purchasing. T and P refer to experimental manipulations (0 = Low, 1 = High). Scale range: 1-7.

Table 6, Mean and Standard Deviation of Initial Openness and Prior Experience by Experimental Cell.

3.8.2 Descriptive statistics and intercorrelations

Table 6 summarizes the descriptive statistics and correlations for all study variables. Overall, participants reported relatively high levels of perceived usefulness (M = 4.48, SD = 1.45), initial openness to AI (M = 4.54, SD = 1.80), trust in the agent (M = 4.15, SD = 1.27), and perceived risk (M = 4.22, SD = 1.41). As expected, purchase intention was positively correlated with trust in the agent (r = .52), perceived usefulness (r = .58), and initial openness (r = .47), and negatively correlated with perceived risk (r = -.47). These correlations are generally consistent with the proposed relationships in the theoretical framework, providing initial support for the model.

Variable	*M*	*SD*	1	2	3	4	5	6	7
1. PI	3.43	1.79							
2. PT	4.34	1.54	.54** [.45, .61]						
3. TIA	4.15	1.27	.52** [.44, .60]	.49** [.40, .57]					
4. PR	4.22	1.41	-.47** [-.55, -.38]	-.30** [-.40, -.20]	-.56** [-.63, -.49]				
5. PU	4.48	1.45	.58** [.51, .65]	.49** [.40, .57]	.74** [.69, .79]	-.49** [-.56, -.40]			
6. AI	3.34	1.66	.70** [.64, .75]	.52** [.44, .59]	.69** [.63, .74]	-.60** [-.66, -.52]	.74** [.69, .79]		
7. IO	4.54	1.80	.47** [.38, .54]	.37** [.27, .46]	.61** [.54, .67]	-.41** [-.49, -.32]	.62** [.55, .68]	.64** [.57, .69]	
8. PRE	2.75	1.58	.42** [.32, .50]	.31** [.21, .40]	.49** [.41, .57]	-.29** [-.38, -.19]	.43** [.34, .51]	.58** [.50, .64]	.62** [.55, .68]

Note. PI = Purchase intention, PT = Perceived transparency, TIA = Trust in the agent, PR = Perceived risk, PU = Perceived usefulness, AI = Adoption intention, IO = Initial openness, PRE = Prior Experience with AI *for Purchasing. M and SD are used to represent mean and standard deviation, respectively. Values in square brackets indicate the 95% confidence interval for each correlation. The confidence interval is a plausible range of population correlations that could have caused the sample correlation (Cumming, 2014). * indicates p < .05. ** indicates p < .01.*

Table 7: means, standard deviations, and correlations with confidence intervals.

3.9 Ethical Conditions

The study was designed in compliance with standard ethical principles for social science research. Participation was entirely voluntary, and participants were free to withdraw at any point without consequence. Informed consent was obtained at the start of the survey: participants were told they were taking part in a study on AI-assisted travel recommendations, though the specific hypotheses were not disclosed in order to avoid demand characteristics (Söderlund, 2018). No personally identifiable information beyond age range and gender was collected, and all data were stored and handled in accordance with GDPR. Participants were fairly compensated through Norstat, in accordance with standard panel research practice. A brief debriefing note at the end of the survey explained the study's purpose and the artificial nature of the stimuli. The study does not involve vulnerable populations, sensitive personal data, or any form of physical or psychological harm; the ethical risks are therefore judged to be minimal.

4. Analysis and Results

4.1 Manipulation Checks

Before testing the main hypotheses, manipulation checks were conducted to verify that the manipulations functioned as intended. Perceived transparency was assessed using a three-item scale ($\alpha = .93$), and perceived price level was assessed using a single item adapted from Lichtenstein et al. (1993).

Transparency manipulation check

Participants in the high-transparency condition reported higher perceived transparency ($M = 4.51$, $SD = 1.53$) than those in the low-transparency condition ($M = 4.17$, $SD = 1.53$), a difference that was statistically significant (diff = 0.34, $t(332) = 2.02$, $p = .044$) but small in magnitude ($d = 0.22$, 95% CI [−0.77, −0.14]; Cohen 1988), raising questions about whether participants experienced the conditions as genuinely distinct. A post-hoc subgroup test indicated that the manipulation produced a moderate difference in perceived transparency in the high-price condition (diff = 0.72, $t(156) = 2.87$, $p = .004$, $d = 0.46$), and no meaningful difference in the low-price condition (diff = 0.01, $t(174) = 0.06$, $p = .953$, $d = 0.01$). A post hoc subgroup test indicated that the manipulation's effect on perceived transparency was significant only in the high-price condition (high price: diff = 0.72, $t(156) = 2.87$, $p = .004$; low price: diff = 0.01, $t(174) = 0.06$, $p = .953$). Combined with the overall weak effect of the manipulation in the main sample, these results indicate that findings on the impact of transparency on trust, perceived risk, perceived usefulness, and purchase intention should be interpreted with caution, as the limited manipulation effect means that observed differences and lack thereof between conditions may not reliably reflect the effect of transparency.

Price manipulation check

Participants in the high-price condition perceived the tickets as significantly more expensive ($M = 4.35$, $SD = 1.48$) than those in the low-price condition ($M = 2.30$, $SD = 1.54$), $t(297) = 11.60$, $p < .001$, indicating a successful manipulation.

4.2 Main Effects

4.2.1 Purchase Intention

4.2.1.1 Main effects analysis

Price

Transparency		High	Low
	High	3.39	3.76
	Low	2.84	3.66

Table 8: Means for purchase intention.

To test hypothesis H1, a two-way ANOVA was conducted with price and transparency as independent variables and purchase intention as the dependent variable. The overall model was significant ($F(3, 337) = 4.60$, $p = .004$), explaining approximately 4% of the variance in purchase intention ($R^2 = .039$). Price had a significant main effect on purchase intention ($F(1, 337) = 9.74$, $p = .002$), with participants in the low-price condition reporting higher purchase intention (M = 3.71) than those in the high-price condition (M = 3.11). Transparency, on the other hand, did not have a significant main effect on purchase intention ($F(1, 337) = 2.98$, $p = .085$), although the effect was in the expected direction, with participants in the high-transparency condition reporting marginally higher purchase intention (M = 3.59) than those in the low-transparency condition (M = 3.28). Finally, the interaction effect between price and transparency was non-significant ($F(1, 337) = 1.39$, $p = .239$), failing to support that price moderates the relationship between transparency and purchase intention.

> Hypothesis H1, which states that transparency has a positive effect on purchase intention, is not supported at the 5% confidence level.

These results, however, may be due to the modest effect of the transparency manipulation or limited statistical power. The non-significant interaction effect is also consistent with the finding that, in the low-price condition, the manipulation failed to create a significant difference in the perceived level of transparency. A non-significant effect of transparency on purchase intention in the low-price condition is therefore to be expected. An exploratory contrast analysis was conducted to examine the effect of transparency separately within each price condition, and to test hypothesis H7, which states that transparency has an effect on purchase intention when price is high.

4.2.1.2 The Effect of Transparency on Purchase Intention under different Price conditions

The results of the contrast analysis showed that in the high-price condition, transparency had a significant positive effect on purchase intention ($F(1, 337) = 3.94$, $p = .048$), with the high-transparency group reporting higher purchase intention (M = 3.39) than the low-transparency group (M = 2.84). In the low-price condition, the effect of transparency on purchase intention was non-significant ($F(1, 337) = 0.16$, $p = .689$).

> Hypothesis H7, which states that transparency has a positive effect on purchase intention when price is high, is supported at the 5% confidence level.

Pursuant to these findings, the transparency causation chain will only be explored under the high price condition, except for H9 (See Chapter 2.8).

4.2.2 Perceived usefulness:

		Price	
		High	Low
Transparency	High	4.60	4.46
	Low	4.28	4.56

Table 9, mean for perceived usefulness.

To test hypothesis H2, a two-way ANOVA was conducted with price and transparency as independent variables and perceived usefulness as the dependent variable. The results revealed that neither price, transparency, nor their interaction had a significant effect on perceived usefulness ($F(3, 348) = 0.82$, $p = .484$, $R^2 = .007$). The cell means (low price / low transparency: M = 4.56; low price / high transparency: M = 4.46; high price / low transparency: M = 4.28; high price / high transparency: M = 4.60) were directionally consistent with the hypothesis, but the effect of transparency was not significant ($F(1, 348)$

= 0.50, p = .478). A contrast analysis confirmed that transparency had no significant effect on perceived usefulness in either the low-price condition ($F(1, 348) = 0.23$, p = .629) or the high-price condition ($F(1, 348) = 2.00$, p = .158). Perceived usefulness is therefore excluded from further analysis.

Hypothesis H2 (and H8) stating that transparency has a positive effect on perceived usefulness (when price is high) is not supported at the 5% confidence level.

4.3 Trust

To test hypotheses H3 and H9, a 2×2 between-subjects ANOVA was conducted. The results showed that transparency had no significant effect on trust ($F(1, 355) = 0.01$, p = .90). The interaction effect between transparency and price was, however, marginally significant ($F(1, 355) = 3.44$, p = .07). While not significant at the conventional 5% level, this result provides directional evidence that the effect of transparency on trust is contingent on price level. A post-hoc exploratory contrast analysis was therefore conducted to provide a more nuanced view of this relationship (Table 10).

		Transparency				
		Low	High	Difference	F	p-value
Price	Low	4.21	3.94	-0.27	2.10	0.148
	High	4.11	4.34	0.23	1.40	0.238
	Joint	-	-	-	1.75	0.175

Table 10, contrast analysis of transparency on trust by price level

While the contrast analysis yielded no significant results, it uncovered a directional finding: among low-price participants, transparency had a negative effect on trust (M-diff = −0.27, p = .148), in contrast to the high-price group, where the effect was positive (M-diff = 0.23, p = .238). This pattern helps explain the near-zero overall effect of transparency on trust in the full sample. While the results lack the significance required to accept hypothesis H9, they provide directional evidence that transparency affects trust differently depending on price level, consistent with the theoretical framework (Chapter 2.8).

Hypothesis H3, stating that transparency has a positive effect on trust, is not supported at the 5% confidence level.

Hypothesis H9, stating that price moderates the relationship between trust and purchase intention, is not supported at the 5% confidence level.

4.4 Trust and Perceived risk as mediators

To test whether the effect of transparency on purchase intention operates through trust and perceived risk (H5 and H6), two mediation analyses were conducted using PROCESS Model 4. Following earlier findings indicating that transparency had no significant effect on purchase intention in the low-price condition, the mediation analyses included only participants in the high-price group (Trust: n = 157).

4.4.1 Trust as a mediator:

The direct effect of transparency on purchase intention was positive but non-significant at the conventional 5% level (effect = 0.402, $t(154) = 1.68$, $p = .095$). The indirect effect through trust was also non-significant (effect = 0.145, 95% CI [−0.164, 0.442]), indicating that trust did not mediate the relationship between transparency and purchase intention. Hypothesis H5 is therefore not supported.

Hypothesis H5, stating that trust mediates the relationship between transparency and purchase intention, is not supported at the 5% confidence level.

4.4.2 Perceived risk as a mediator:

As with trust, perceived risk did not significantly mediate the relationship between transparency and purchase intention. The direct effect was positive and significant (effect

= 0.560, $t(154) = 2.15$, $p = .033$). The indirect effect, however, was not only non-significant but also close to zero (effect = −0.007, 95% CI [−0.129, 0.118]), a finding that does not align with the theoretical framework (Chapter 2). These results indicate that perceived risk does not mediate the relationship between transparency and purchase intention. Hypothesis H6 is therefore rejected.

> Hypothesis H6, stating that perceived risk mediates the relationship between transparency and purchase intention, is rejected.

4.4.3 Summary of the mediation analyses:

Neither trust nor perceived risk mediated the relationship between transparency and purchase intention. For perceived risk specifically, the indirect effect was close to zero, a finding inconsistent with hypothesis H6, which is therefore rejected. For trust, while the results were directionally consistent with the hypothesis, they were not significant; the hypothesis is therefore not supported. Limiting the analysis to the high-price group played a role in weakening the results in two key ways: (1) it reduced the sample size to 157 participants, which is underpowered for a mediation analysis; (2) most of the variation in perceived financial risk could be attributed to price, so restricting the sample to the high-price group reduced the variation in the perceived risk variable.

4.5 Effect of the control variables

To assess the robustness of the main findings, the analyses presented in Sections 4.2–4.4 were re-estimated, including initial openness, prior experience, age, and gender as control variables. The results of these analyses are reported below and compared with the original models.

4.5.1 Controlled effect of price and transparency on purchase intention:

The results differed from the original analysis. Price remained significant ($F = 11.86$, $p < .001$), while transparency ($F = 0.43$, $p = .511$) and the interaction between price and transparency ($F = 0.50$, $p = .480$) were not significant. Compared to the initial results, the transparency effect was substantially reduced.

The exploratory contrast analysis was also re-estimated on the controlled model. Contrary to the original analysis, transparency did not have a significant effect in the high-price condition ($F = 0.86$, $p = .355$). The effect also remained non-significant in the low-price condition ($F = 0.00$, $p = .975$).

4.5.2 Controlled effect of on trust under different price conditions:

For the controlled trust analysis, neither transparency ($F = 1.17$, $p = .281$) nor the interaction between price and transparency ($F = 0.04$, $p = .834$) were significant. Compared to the original model, the interaction effect became substantially weaker after the control variables were included. Initial openness ($F = 12.91$, $p < .001$) and prior experience ($F = 2.18$, $p = .045$) were significant predictors of trust, whereas age ($F = 0.43$, $p = .984$) and gender ($F = 0.54$, $p = .464$) were not.

The exploratory contrast analysis was also conducted on the controlled model. Transparency did not have a significant effect in either the low-price condition ($F = 0.91$, $p = .341$) or the high-price condition ($F = 0.35$, $p = .555$). The joint test was likewise non-significant ($F = 0.63$, $p = .534$). While the original model suggested a negative effect of transparency on trust in the low-price condition and a positive effect in the high-price condition, this pattern became substantially weaker after the control variables were included.

4.5.3 Controlled mediation analysis: Trust

Trust remained a significant predictor of purchase intention ($b = 0.52$, $p < .001$), while transparency did not have a significant effect on trust ($b = -0.18$, $p = .263$). Among the control variables, both initial openness ($b = 0.37$, $p < .001$) and prior experience ($b = 0.17$, $p = .009$) were positively associated with trust. Prior experience was also positively associated with purchase intention ($b = 0.38$, $p < .001$).

Despite these relationships, the indirect effect of transparency on purchase intention through trust was not significant (effect = -0.093, 95% CI [-0.286, 0.049]). As in the initial analysis, the confidence interval included zero, indicating that trust did not mediate the relationship between transparency and purchase intention.

4.5.4 Controlled mediation analysis: Perceived risk

The results were largely unchanged from the original analysis. Transparency did not have a significant effect on perceived risk ($b = 0.18$, $p = .391$), while perceived risk remained a significant predictor of purchase intention ($b = -0.37$, $p < .001$).
The indirect effect of transparency on purchase intention through perceived risk was not significant (effect = -0.067, 95% CI [-0.245, 0.089]). As in the original analysis, the confidence interval included zero, indicating that perceived risk did not mediate the relationship between transparency and purchase intention.

4.5.3 Summary of the controlled analyses:

The inclusion of initial openness, prior experience, age, and gender affected some of the findings presented in the original analyses. In the purchase intention analysis, the previously significant effect of transparency in the high-price condition was no longer significant after the control variables were included. Similarly, the directional pattern observed in the trust analysis became weaker.

The mediation analyses, however, were largely unchanged. Neither trust nor perceived risk mediated the relationship between transparency and purchase intention, and both indirect effects remained non-significant after the control variables were included.

It is important to note that initial openness and prior experience were measured after exposure to the experimental stimulus and varied significantly across experimental conditions, particularly in the high-price/high-transparency condition (see Chapter 3.8.1.2). Consequently, the controlled analyses should be interpreted with caution.

4.6 Summary of hypothesis outcomes

The effect of transparency and price on both dependent variables was first tested using a 2×2 between-subjects ANOVA. For purchase intention, only the effect of price was significant at the 5% confidence level. The effect of transparency was marginally significant, with a *p*-value of .085.

For perceived usefulness, neither the independent variables nor the interaction term had a significant effect. Hypothesis H2 (and H8) was therefore not supported. For purchase intention, while the interaction term was not significant, the effect of transparency was marginally significant, and the effect of price was significant. The follow-up contrast analysis showed that the effect of transparency on purchase intention was significant only when price was high. Hypothesis H7 was therefore supported. However, because the interaction effect was not significant, this finding should be interpreted as suggestive rather than conclusive evidence of a conditional effect. The inclusion of the control variables did not alter these conclusions, as neither trust nor perceived risk emerged as a significant mediator in the controlled analyses.

The results of both mediation analyses, through trust and through perceived risk, were not significant. The findings for the hypothesis that trust mediated the relationship between transparency and purchase intention were directionally consistent with the theoretical framework but lacked the required significance level, leaving the related hypothesis unsupported. The findings for the hypothesis that perceived risk mediated the relationship between transparency and purchase intention were not significant and not consistent with the theoretical framework, leading to the rejection of the related hypothesis. The inclusion of the control variables did not alter these conclusions, as neither trust nor perceived risk emerged as significant mediators in the controlled analyses.

Hypothesis	Statement	Type
H1	Higher transparency leads to higher purchase intention.	Not supported
H2	Higher transparency leads to higher perceived usefulness.	Not supported
H3	Higher transparency leads to higher trust.	Not Supported
H4	The relationship between transparency and purchase intention is mediated by trust.	Not supported
H5	The relationship between transparency and perceived usefulness is mediated by trust.	Not Supported
H6	The relationship between transparency and purchase intention is mediated by perceived risk.	Rejected
H7	Higher transparency leads to higher purchase intention when price is high.	Supported
H8	Higher transparency leads to higher perceived usefulness when price is high.	Not supported
H9	The relationship between transparency and trust is moderated by price.	Not supported

Table 11, Summary of tested Hypotheses outcome

5. Discussion

This chapter interprets the empirical findings of Chapter 4 together with the theoretical framework developed in Chapter 2. The discussion proceeds in four parts. Section 5.1 addresses the theoretical implications for the literature on AI transparency, perceived risk, and consumer trust in recommender systems. Section 5.2 translates these findings into managerial implications for airlines and online travel agencies deploying AI-assisted booking tools. Section 5.3 presents the limitations of the study and directions for future research.

5.1 Theoretical Implications

The central contribution of this thesis is suggestive evidence that the effect of AI transparency on consumer behaviour is conditional. Four key contributions can be presented:

First, the recommender systems literature has typically presented transparency as having a positive effect on purchase intentions and purchasing behaviour (see Chapter 2.3). This study explores price as a potential boundary condition. Although the interaction between price and transparency was not statistically significant, the analyses presented in Chapter 4.2 revealed that the positive relationship between transparency and purchase intention was observed only in the high-price condition. However, this effect was no longer significant when initial openness, prior experience, age, and gender were included as control variables. Given the weak transparency manipulation, the absence of a significant interaction effect, and the sensitivity of the finding to the inclusion of control variables, the results should be interpreted as suggestive rather than definitive evidence that transparency may be more effective in high-price contexts.

Second, while there are guidelines for what constitutes a good explanation in general and in the context of recommender systems (see Chapter 2.3), these guidelines do not account for the importance of, or need for, an explanation or how it should be adapted across contexts. This study shows that despite having a structurally identical explanation across both price conditions, the resulting level of perceived transparency differed between the groups (see Chapter 4.1). One potential conclusion that could be drawn from such a finding is that the guidelines for what constitutes a good explanation should vary across price levels. It is also important to consider, in interpreting these findings, that the displayed itineraries across the two price conditions differed substantially in complexity (high price: Stockholm – Rio de Janeiro; low price: Stockholm – Warsaw). (See Chapter 3.x)

Third, while the effect of transparency on trust was not significant at the conventional 5% level, the results suggest that the effect of transparency on trust is not always positive. This finding is in line with Kizilcec (2016) and Scharowski et al. (2022), who identify contexts in which the positive effect of explainability is reversed. Kizilcec (2016) provides one potential explanation for the negative effect of transparency on trust in low-price contexts: the explanation text may have been too lengthy and led to unnecessary cognitive overload.

Finally, perceived risk and trust were conceptualized as the psychological mechanisms through which transparency would influence purchase intention. This choice was grounded in two complementary streams of work. Pavlou's e-commerce model identifies trust and perceived risk, alongside perceived usefulness and perceived ease of use, as central determinants of consumers' intention to transact under uncertainty, and Cox and Rich's (1964) account, around which the present study was designed, similarly highlights the uncertainty inherent in high-stakes purchase situations. In a high-price context, the amount at stake is larger and the need for certainty correspondingly greater; if this need is not met, perceived risk remains high, and the purchase is unlikely to be made. Pavlou's model reflects a similar logic by positioning trust and perceived risk as central mechanisms for managing such uncertainty.

In the present study, however, these mechanisms did not operate as originally anticipated. Transparency did not influence purchase intention via perceived risk or trust in the AI assistant in the way the mediation hypotheses implied. Within the tested models, information in the form of an explanation did not significantly reduce perceived risk, nor did it meaningfully increase trust in the agent if this is interpreted as a coarse proxy for certainty. At the same time, transparency did significantly affect purchase intention in the high-price condition. Together with the findings reported in Chapters 4.1 and 4.2, this pattern suggests that transparency had a behavioural impact, but not through the global trust and risk pathways that were specified a priori.

For perceived risk, one explanation is that the scale used in this study primarily captured financial risk. Given this focus, the moderate effect size of the transparency manipulation in the high-price condition may simply have been too small to substantially lower perceived financial risk. Moreover, when interpreted through Cox and Rich's logic, transparency is expected to reduce perceived risk indirectly by increasing certainty that the purchase will have favourable consequences, rather than exerting a strong direct effect on risk ratings. In Cox and Rich's framework, certainty is shaped not only by new information but also by prior experiences and the experiences of others, which were not fully captured in the present design. This limits the extent to which a single transparency intervention can be expected to change perceived risk scores.

For trust, the picture is more complex. Even if transparency can influence trust in an AI system, a one-off explanation may not be sufficient to counteract existing attitudes and expectations about AI. Models of initial trust in e-commerce, such as those developed by McKnight and colleagues, as well as Pavlou's work, explicitly include structural assurances, reputation, and prior experience as determinants of baseline trust. In the AI domain, prior attitudes and initial openness to AI are likely to be even more salient. Many people hold strong views about AI's reliability, fairness, and potential harms, and these predispositions may shape how willing they are to accept vulnerability to an AI system. Users who fundamentally distrust AI may remain skeptical even when presented with an explanation, whereas users who already have a positive attitude towards AI may follow its recommendations without needing additional transparency. Under these conditions, a moderate transparency manipulation is unlikely to produce large changes in global agent-level trust.

Taken together, these results suggest that the primary mechanism behind the transparency effect in the high-price scenario is more specific than global trust in the agent or general perceived risk. A more precise candidate is a construct that captures certainty that following the recommendation will lead to a favourable outcome. Trust in the agent contributes to this sense of certainty but is not identical to it. Conceptually, such a mechanism would be located at the level of the recommendation itself and could be expressed in terms of perceived accuracy, credibility, or trust in the recommendation. This recommendation-level evaluation is influenced by, but distinct from, agent-level trust, and it aligns more closely with Cox and Rich's notion of reducing uncertainty about a particular purchase decision. Under this view, transparency may first shape how credible and reliable the specific recommendation appears and, through that, affect recommendation-level trust, perceived risk, and ultimately purchase intention, rather than directly altering global agent-level trust or general risk perceptions. The finding that transparency significantly influenced purchase intention in the high-price condition, while perceived risk and trust in the agent did not mediate this effect, supports the idea that the crucial mechanism is located at the level of the specific recommendation rather than in global trust or perceived risk alone.

5.2 Managerial Implications

The results of this study carry three key findings for practitioners in the air travel sector using AI-powered flight recommender systems:

First, for complex and expensive itineraries, transparency can play a significant role in shifting the customer's willingness to buy. Such explanations should be based on the findings of Miller (2019) and communicate to the buyer, in concrete terms, why the recommended option fits their criteria best compared to all other options.

Second, what constitutes an effective explanation in a high-price context may differ from what works in a low-price context. The findings of this study indicate that, despite following the same design principles across both price conditions, the level of perceived transparency differed substantially. This points to the need to design explanations differently depending on price level.

Third, while non-significant, the effect of transparency on trust in the low-price context was negative. Caution should therefore be exercised when implementing explainability in a recommender system context. The content of the explanation should be drafted carefully; both accuracy and conciseness must be maintained. Without accuracy, the user may perceive the system as less reliable, and an overly lengthy explanation may lead to cognitive overload and, by extension, a loss of trust. This finding remains directional and non-significant and should therefore be interpreted with caution.

Finally, while the results indicate that transparency leads to higher purchase intention in a high-price context, they failed to find support for perceived usefulness, which is a key determinant of adoption intention according to the TAM. This null finding suggests that, given the presence of advanced and well-established flight booking platforms, a system similar to the one designed for this study may not gain commercial traction among the studied population (Swedish adults aged 20–39). This claim, however, is an interpretation of the results and should be read as an indication rather than a firm implication.

5.3 Limitations

While this study provides valuable theoretical and managerial contributions, several key limitations should be considered.

First, while the transparency manipulation was significant overall, its effect on perceived transparency was weak and differed substantially across the two price conditions (see Chapters 2.3, 4.1, and 5.1). Findings related to the low-price context should therefore be interpreted carefully and considering the transparency manipulation itself (see Appendix A). Future research on explanation design should explore the elements needed to achieve effective levels of explainability under different levels of stakes and complexity.

Second, the context of the study itself (flight booking recommendations) implied that the two recommended itineraries differed not only in price but also in complexity (see Chapter 3.1). Complexity plays a role in the usefulness and necessity of the explanation text, which was not accounted for in the design of the survey. Future research could replicate this study while controlling for such confounds.

Third, due to the design of the study, purchase intention was used as a proxy for purchasing behaviour. While the relationship between the two is well-established, the two concepts remain distinct (see Chapter 2.2.1). Future research could replicate this study in a real-world setting to assess the true effect of perceived risk on purchasing behaviour.

Fourth, the sample is restricted to Swedish adults between the ages of 20 and 39. The purpose of this restriction is to reduce variance from the demographic moderators specified in UTAUT (Venkatesh et al., 2003) and from cultural differences across geographies, given the limited sample size (see Chapter 1.4). Future research could conduct a similar experiment with a broader age range and in a different cultural context.

Fifth, this study employed shortened measurement scales for many key variables, including the main dependent variable (purchase intention), to minimize survey length and the resulting fatigue and risk of speeding and straightlining. Shortened and single-item scales have been shown to be less reliable, and removing items from validated scales may reduce validity (see Chapter 3.3).

Sixth, initial openness and prior experience both play a key role in the context of AI and technology adoption (see Chapter 2.7). These variables were measured after participants had been exposed to the experimental stimulus. As a result, they may not solely reflect pre-existing individual differences but may also have been influenced by participants' reactions to the treatment. This interpretation is supported by the finding that both initial openness and prior experience differed significantly across experimental conditions. (see Chapter 3.8.1.2) While robustness checks, including these variables as controls, altered some of the findings, particularly those related to purchase intention, their inclusion may have adjusted for treatment-related variance in addition to baseline characteristics. Future research could measure these variables prior to treatment exposure and incorporate them into the analysis to more clearly distinguish between pre-existing individual differences and responses to the experimental stimulus.

Finally, by limiting the mediation analysis to the high-price group, the sample size dropped to N = 157, which limits the statistical power to detect indirect effects.

6. Conclusion

This study returns to the puzzle raised in Chapter 1.2: when an AI recommends a flight, does explaining the recommendation make the consumer more likely to act on it?

RQ1: Does transparency affect purchase intention?

The findings of this study indicate that transparency does not have a significant effect on purchase intention. The main effect of transparency on purchase intention was not significant in the two-way ANOVA. The transparency manipulation, however was weak, and these results should be interpreted with caution.

RQ2: Does transparency have a stronger effect on purchase intention when price is high?

The interaction effect between price and transparency on purchase intention in the two-way ANOVA was not significant. However, a contrast analysis revealed that transparency had a significant effect only in the high-price condition. These results remain inconclusive, as the manipulation failed to create a significant difference in transparency perception within the low-price group. Furthermore, the effect observed in the high-price condition was no longer significant when initial openness, prior experience, age, and gender were included as control variables, suggesting that this finding should be interpreted with caution.

RQ3: Through which mechanism (trust, perceived risk) does transparency affect purchase intention?

Neither of the identified psychological mechanisms (trust and perceived risk) mediated the relationship between transparency and purchase intention. It is important to note, however, that emphasis was placed on financial risk when developing the scale, and the results are therefore mainly applicable to financial risk. The combination of the moderate effect of the manipulation on perceived transparency and the relatively low sample size might, however explain the null findings, especially for trust as a mediator. Additionally, the theoretical framework adopted in this case, in addition to the Cox and Rich (1964), would point towards trust as a mediator between transparency and perceived risk, a relationship not tested in this study.

Closing remarks

The positive effect of transparency on purchasing behavior is well-established in the recommender systems literature. This study contributes to the understanding of this relationship by exploring price as a potential boundary condition. While the study has limitations, it sheds light on two key findings: (1) transparency has a significant positive effect on purchase intention when price is high, although this finding was not robust to the inclusion of control variables and should therefore be interpreted with caution; (2) what is considered an effective explanation in the context of recommender systems differs based on the price and complexity of the product.

References

Afroogh, S., Akbari, A., Malone, E. *et al.* Trust in AI: progress, challenges, and future directions. *Humanit Soc Sci Commun* **11**, 1568 (2024). https://doi.org/10.1057/s41599-024-04044-8

Ajzen, I. (1991). The theory of planned behavior. Organizational Behavior and Human Decision Processes, 50(2), 179–211. https://doi.org/10.1016/0749-5978(91)90020-T

Animesh, A., Pinsonneault, A., Yang, S. B., & Oh, W. (2011). An odyssey into virtual worlds: Exploring the impacts of technological and spatial environments on intention to purchase virtual products. MIS Quarterly, 35(3), 789–810. https://www.misq.org/an-odyssey-into-virtual-worlds-exploring-the-impacts-of-technological-and-spatial-environments-on-intention-to-purchase-virtual-products.html

Barredo Arrieta, A., Díaz-Rodríguez, N., Del Ser, J., Bennetot, A., Tabik, S., Barbado, A., García, S., Gil-Lopez, S., Molina, D., Benjamins, R., Chatila, R., & Herrera, F. (2020). Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI. Information Fusion, 58, 82–115. https://doi.org/10.1016/j.inffus.2019.12.012

Bauer, R. A. (1960). Consumer behavior as risk taking. In R. S. Hancock (Ed.), Dynamic marketing for a changing world: Proceedings of the 43rd National Conference of the American Marketing Association (pp. 389–398). American Marketing Association.

Bergkvist, L., & Rossiter, J. R. (2007). The predictive validity of multiple-item versus single-item measures of the same constructs. Journal of Marketing Research, 44(2), 175–184. https://doi.org/10.1509/jmkr.44.2.175

Broström, A., & Wennberg, K. (2025). Data, models, and insights: A handbook for analysis. KTH & Stockholm School of Economics.

Bryman, A., & Bell, E. (2011). Business research methods (3rd ed.). Oxford University Press.

Chen, C., Tian, A. D., & Jiang, R. (2024). When post hoc explanation knocks: Consumer responses to explainable AI recommendations. Journal of Interactive Marketing, 59(3), 234–250. https://doi.org/10.1177/10949968231200221

Churchill, G. A., Jr. (1979). A paradigm for developing better measures of marketing constructs. Journal of Marketing Research, 16(1), 64–73. https://doi.org/10.1177/002224377901600107

Cohen, J. (1992). A power primer. Psychological Bulletin, 112(1), 155–159. https://doi.org/10.1037/0033-2909.112.1.155

Cook, T. D., & Campbell, D. T. (1979). Quasi-experimentation: Design and analysis issues for field settings. Houghton Mifflin.

Cox, D. F., & Rich, S. U. (1964). Perceived risk and consumer decision-making: The case of telephone shopping. Journal of Marketing Research, 1(4), 32–39. https://doi.org/10.2307/3150375

Cramer, H., Evers, V., Ramlal, S., van Someren, M., Rutledge, L., Stash, N., Aroyo, L., & Wielinga, B. (2008). The effects of transparency on trust in and acceptance of a content-based art recommender. User Modeling and User-Adapted Interaction, 18(5), 455–496. https://doi.org/10.1007/s11257-008-9051-3

Cumming, G. (2014). The new statistics: Why and how. Psychological Science, 25(1), 7–29. https://doi.org/10.1177/0956797613504966

Dabholkar, P. A., & Sheng, X. (2012). Consumer participation in using online recommendation agents: Effects on satisfaction, trust, and purchase intentions. The Service Industries Journal, 32(9), 1433–1449. https://doi.org/10.1080/02642069.2011.624596

Dadoun, A., Defoin-Platel, M., Fiig, T., Landra, C., & Troncy, R. (2021). How recommender systems can transform airline offer construction and retailing. Journal of Revenue and Pricing Management, 20(3), 301–315.

Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. MIS Quarterly, 13(3), 319–340. https://doi.org/10.2307/249008

Davis, F. D., Bagozzi, R. P., & Warshaw, P. R. (1989). User acceptance of computer technology: A comparison of two theoretical models. Management Science, 35(8), 982–1003. https://doi.org/10.1287/mnsc.35.8.982

Eisinga, R., Grotenhuis, M., & Pelzer, B. (2013). The reliability of a two-item scale: Pearson, Cronbach, or Spearman-Brown? International Journal of Public Health, 58(4), 637–642. https://doi.org/10.1007/s00038-012-0416-3

Eurostat. (2026). Digital skills in 2023: Impact of education and age [News article]. European Commission. https://ec.europa.eu/eurostat/web/products-eurostat-news/w/ddn-20240222-1

Fishbein, M., & Ajzen, I. (1975). Belief, attitude, intention, and behavior: An introduction to theory and research. Addison-Wesley.

Fox, J. (2007). The uncertain relationship between transparency and accountability. Development in Practice, 17(4–5), 663–671. https://doi.org/10.1080/09614520701469955

Gefen, D., Karahanna, E., & Straub, D. W. (2003). Trust and TAM in online shopping: An integrated model. MIS Quarterly, 27(1), 51–90. https://doi.org/10.2307/30036519

Gillespie, N., Lockey, S., & Curtis, C. (2025). Trust, attitudes and use of artificial intelligence: A global study 2025. University of Melbourne and KPMG.

Glikson, E., & Woolley, A. W. (2020). Human trust in artificial intelligence: Review of empirical research. Academy of Management Annals, 14(2), 627–660. https://doi.org/10.5465/annals.2018.0057

Gomez-Uribe, C. A., & Hunt, N. (2016). The Netflix recommender system: Algorithms, business value, and innovation. ACM Transactions on Management Information Systems, 6(4), 1–19. https://doi.org/10.1145/2843948

Greszki, R., Meyer, M., & Schoen, H. (2015). Exploring the effects of removing "too fast" responses and respondents in online surveys. Public Opinion Quarterly, 79(2), 471–503. https://doi.org/10.1093/poq/nfu058

Gursoy, D., Chi, O. H., Lu, L., & Nunkoo, R. (2019). Consumers acceptance of artificially intelligent (AI) device use in service delivery. International Journal of Information Management, 49, 157–169. https://doi.org/10.1016/j.ijinfomgt.2019.03.008

Häubl, G., & Trifts, V. (2000). Consumer decision making in online shopping environments: The effects of interactive decision aids. Marketing Science, 19(1), 4–21. https://doi.org/10.1287/mksc.19.1.4.15178

Jacoby, J., & Kaplan, L. B. (1972). The components of perceived risk. In M. Venkatesan (Ed.), Proceedings of the Third Annual Conference of the Association for Consumer Research (pp. 382–393). Association for Consumer Research.

Jarvenpaa, S. L., Tractinsky, N., & Vitale, M. (2000). Consumer trust in an Internet store. Information Technology and Management, 1(1–2), 45–71. https://doi.org/10.1023/A:1019104520776

Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. Econometrica, 47(2), 263–291. https://doi.org/10.2307/1914185

Kan, H. Y., Wong, D., & Chau, K. (2024). A personalized flight recommender system based on link prediction in aviation data. IEEE Access, 12. https://doi.org/10.1109/ACCESS.2024.3369487

Kaushik, A. K., Agrawal, A. K., & Rahman, Z. (2015). Tourist behaviour towards self-service hotel technology adoption: Trust and subjective norm as key antecedents. Tourism Management Perspectives, 16, 278–289. https://doi.org/10.1016/j.tmp.2015.09.002

Kim, M. J., Lee, C.-K., & Jung, T. (2020). Exploring consumer behavior in virtual reality tourism using an extended stimulus-organism-response model. Journal of Travel Research, 59(1), 69–89. https://doi.org/10.1177/0047287518818915

Kim, Y., Dykema, J., Stevenson, J., Black, P., & Moberg, D. P. (2019). Straightlining: Overview of measurement, comparison of indicators, and effects in mail-web mixed-mode surveys. Social Science Computer Review, 37(2), 214–233. https://doi.org/10.1177/0894439317752406

Kizilcec, R. F. (2016). How much information? Effects of transparency on trust in an algorithmic interface. In Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems (pp. 2390–2395). Association for Computing Machinery. https://doi.org/10.1145/2858036.2858402

Lichtenstein, D. R., Ridgway, N. M., & Netemeyer, R. G. (1993). Price perceptions and consumer shopping behavior: A field study. Journal of Marketing Research, 30(2), 234–245. https://doi.org/10.2307/3172830

Lim, C., & Kim, S. (2024). Examining factors influencing the user's loyalty on algorithmic news recommendation service. Humanities and Social Sciences Communications, 11, Article 91. https://doi.org/10.1057/s41599-023-02516-x

Linden, G., Smith, B., & York, J. (2003). Amazon.com recommendations: Item-to-item collaborative filtering. IEEE Internet Computing, 7(1), 76–80. https://doi.org/10.1109/MIC.2003.1167344

Ma, X., & Huo, Y. (2023). Are users willing to embrace ChatGPT? Exploring the factors on the acceptance of chatbots from the perspective of AIDUA framework. Technology in Society, 75, 102362. https://doi.org/10.1016/j.techsoc.2023.102362

Mayer, R. C., Davis, J. H., & Schoorman, F. D. (1995). An integrative model of organizational trust. Academy of Management Review, 20(3), 709–734. https://doi.org/10.5465/amr.1995.9508080335

McKnight, D. H. (2011). Trust in information technology. In G. B. Davis (Ed.), The Blackwell Encyclopedia of Management: Management Information Systems (2nd ed., pp. 1–6). Wiley-Blackwell.

McKnight, D. H., Choudhury, V., & Kacmar, C. (2002). Developing and validating trust measures for e-commerce: An integrative typology. Information Systems Research, 13(3), 334–359. https://doi.org/10.1287/isre.13.3.334.81

McKnight, D. H., Cummings, L. L., & Chervany, N. L. (1998). Initial trust formation in new organizational relationships. Academy of Management Review, 23(3), 473–490. https://doi.org/10.5465/amr.1998.926622

Mehrabian, A., & Russell, J. A. (1974). An approach to environmental psychology. MIT Press.

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

Mitchell, V.-W. (1999). Consumer perceived risk: Conceptualisations and models. European Journal of Marketing, 33(1/2), 163–195. https://doi.org/10.1108/03090569910249229

Ngo, V. M. (2025). Balancing AI transparency: Trust, certainty, and adoption. Information Development. Advance online publication. https://doi.org/10.1177/02666669251346124

Nunnally, J. C. (1978). Psychometric theory (2nd ed.). McGraw-Hill.

Park, K., & Yoon, H. Y. (2025). AI algorithm transparency, pipelines for trust not prisms: Mitigating general negative attitudes and enhancing trust toward AI. Humanities and Social Sciences Communications, 12, Article 1160. https://doi.org/10.1057/s41599-025-05116-z

Pavlou, P. A. (2003). Consumer acceptance of electronic commerce: Integrating trust and risk with the Technology Acceptance Model. International Journal of Electronic Commerce, 7(3), 101–134. https://doi.org/10.1080/10864415.2003.11044275

Petty, R. E., & Cacioppo, J. T. (1986). Communication and persuasion: Central and peripheral routes to attitude change. Springer-Verlag. https://doi.org/10.1007/978-1-4612-4964-1

Pew Research Center. (2025, October 15). How people around the world view AI. https://www.pewresearch.org/global/2025/10/15/how-people-around-the-world-view-ai/

Preacher, K. J., & Hayes, A. F. (2008). Asymptotic and resampling strategies for assessing and comparing indirect effects in multiple mediator models. Behavior Research Methods, 40(3), 879–891. https://doi.org/10.3758/BRM.40.3.879

Pu, P., Chen, L., & Hu, R. (2011). A user-centric evaluation framework for recommender systems. Proceedings of the Fifth ACM Conference on Recommender Systems (RecSys '11), 157–164.

Qualtrics. (2024, August 15). *Response quality*. https://www.qualtrics.com/support/survey-platform/survey-module/survey-checker/response-quality/

Rahman, W. (2025). Explainable AI–driven recommender systems in e-commerce: Effects on user trust, perceived transparency, and purchase decisions [Preprint]. SSRN. https://doi.org/10.2139/ssrn.5789189

Rosenberger, J., Kuhlemann, S., Tiefenbeck, V., Kraus, M., & Zschech, P. (2025). The impact of transparency in AI systems on users' data-sharing intentions: A scenario-based experiment. arXiv. https://arxiv.org/abs/2502.20243

Scharowski, N., Perrig, S. A. C., von Felten, N., & Brühlmann, F. (2022). Trust and reliance in XAI: Distinguishing between attitudinal and behavioral measures. arXiv. https://arxiv.org/abs/2203.12318

Schmidt, P., Biessmann, F., & Teubner, T. (2020). Transparency and trust in artificial intelligence systems. Journal of Decision Systems, 29(4), 260–278. https://doi.org/10.1080/12460125.2020.1819094

Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). Experimental and quasi-experimental designs for generalized causal inference. Houghton Mifflin.

Sheeran, P. (2002). Intention–behavior relations: A conceptual and empirical review. European Review of Social Psychology, 12(1), 1–36. https://doi.org/10.1080/14792772143000003

Shi, S., Gong, Y., & Gursoy, D. (2021). Antecedents of trust and adoption intention toward artificially intelligent recommendation systems in travel planning: A heuristic–systematic model. Journal of Travel Research, 60(8), 1714–1734.

Sinha, R., & Swearingen, K. (2002). The role of transparency in recommender systems. In Proceedings of the CHI 2002 Conference on Human Factors in Computing Systems: Extended Abstracts (pp. 830–831). ACM. https://doi.org/10.1145/506443.506619

Slade, E. L., Dwivedi, Y. K., Piercy, N. C., & Williams, M. D. (2015). Modeling consumers' adoption intentions of remote mobile payments in the United Kingdom: Extending UTAUT with innovativeness, risk, and trust. Psychology & Marketing, 32(8), 860–873. https://doi.org/10.1002/mar.20823

Söderlund, M. (2018). Experiments in marketing. Studentlitteratur.

Söderlund, M., Vilgon, M., & Gunnarsson, J. (2001). Predicting purchasing behavior on business-to-business markets. European Journal of Marketing, 35(1/2), 168–181. https://doi.org/10.1108/03090560110363409

Statista. (2026). Flight search engine online bookings by brand in Sweden as of June 2024. Statista Inc. https://www.statista.com/forecasts/1348175/flight-search-engine-online-bookings-by-brand-in-sweden

Tintarev, N., & Masthoff, J. (2007). A survey of explanations in recommender systems. In Proceedings of the 2007 IEEE 23rd International Conference on Data Engineering Workshop (pp. 801–806). IEEE. https://doi.org/10.1109/ICDEW.2007.4401070

Venkatesh, V., & Bala, H. (2008). Technology acceptance model 3 and a research agenda on interventions. Decision Sciences, 39(2), 273–315. https://doi.org/10.1111/j.1540-5915.2008.00192.x

Venkatesh, V., & Davis, F. D. (2000). A theoretical extension of the technology acceptance model: Four longitudinal field studies. Management Science, 46(2), 186–204. https://doi.org/10.1287/mnsc.46.2.186.11926

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

Venkatesh, V., Thong, J. Y. L., & Xu, X. (2012). Consumer acceptance and use of information technology: Extending the unified theory of acceptance and use of technology. MIS Quarterly, 36(1), 157–178. https://doi.org/10.2307/41410412

Völckner, F. (2008). The dual role of price: Decomposing consumers' reactions to price. Journal of the Academy of Marketing Science, 36(3), 359–377. https://doi.org/10.1007/s11747-007-0076-7

Wang, W., & Benbasat, I. (2009). Interactive decision aids for consumer decision making in e-commerce: The influence of perceived strategy restrictiveness. MIS Quarterly, 33(2), 293–320. https://doi.org/10.2307/20650293

Wang, X., & Yin, M. (2021). Are explanations helpful? A comparative study of the effects of explanations in AI-assisted decision making. In Proceedings of the 26th International Conference on Intelligent User Interfaces (pp. 318–328). ACM. https://doi.org/10.1145/3397481.3450650

Wanner, J., Herm, L.-V., Heinrich, K., & Janiesch, C. (2022). The effect of transparency and trust on intelligent system acceptance: Evidence from a user-based study. Electronic Markets, 32(4), 2079–2102. https://doi.org/10.1007/s12525-022-00593-5

Warshaw, P. R., & Davis, F. D. (1985). Disentangling behavioral intention and behavioral expectation. Journal of Experimental Social Psychology, 21(3), 213–228. https://doi.org/10.1016/0022-1031(85)90017-4

Werz, J. M., Borowski, E., & Isenhardt, I. (2024). Explainability as a means for transparency? Lay users' requirements towards transparent AI. In L. Paletta (Ed.), *Cognitive Computing and Internet of Things* (Vol. 124). AHFE International. https://doi.org/10.54941/ahfe1004712

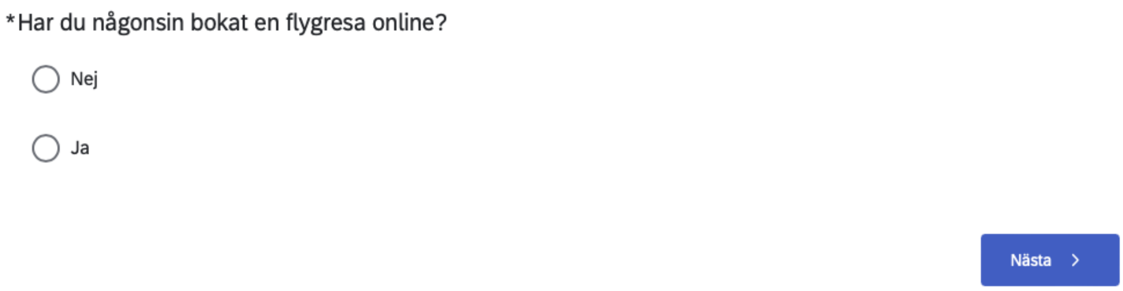
Zanpure, S. (2025, December 23). *Online travel is set for growth, but AI threatens the status quo.* Skift. https://skift.com/2025/12/23/online-travel-is-set-for-growth-but-ai-threatens-the-status-quo/

Appendix

Appendix A: Survey

The following pages present the full survey instrument as administered to participants via Norstat. The survey was presented in Swedish. The four experimental conditions are shown in sequence below.

Appendix X, our survey


*Har du någonsin bokat en flygresa online?

Nej

Ja

Nästa

*Ålder

"Hur gammal är du?"

*Kön

- Man
- Kvinna
- Icke-binär
- Föredrar att inte ange
- Annat

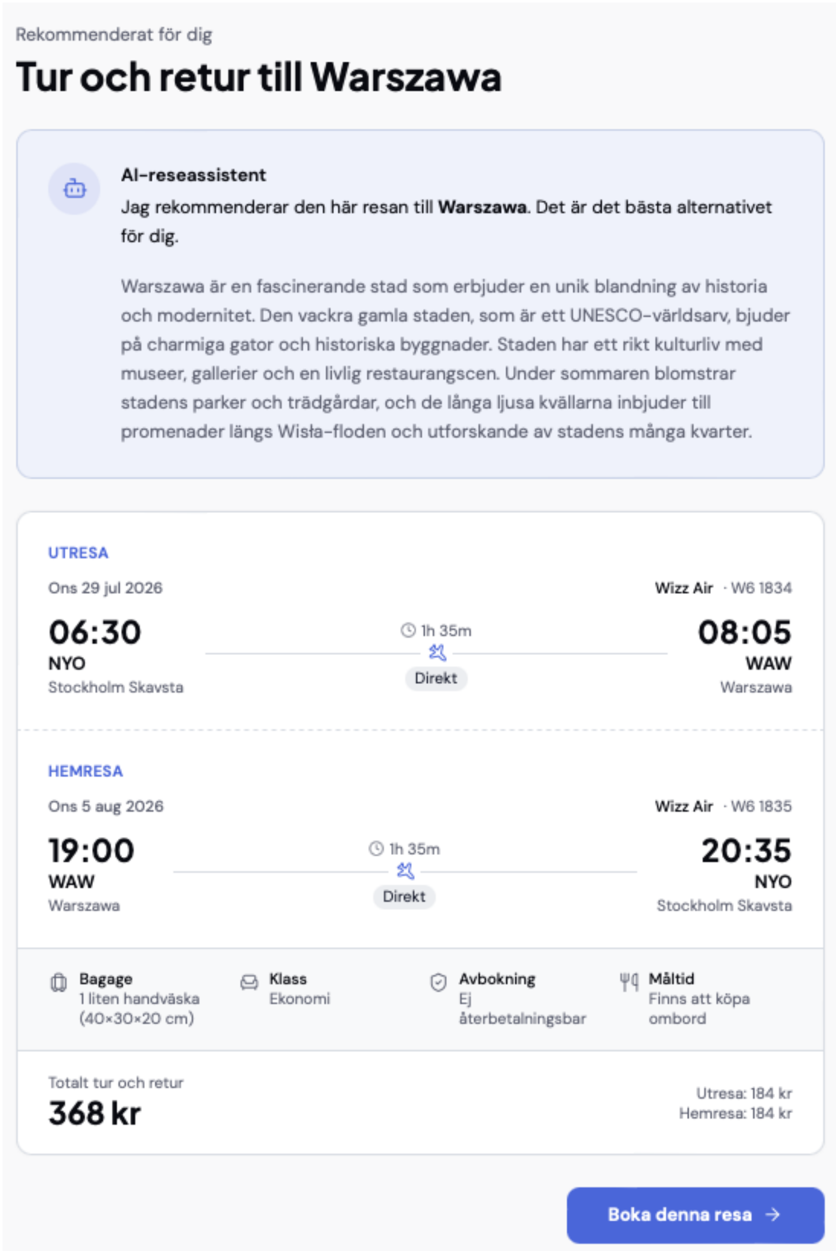
Nästa >

After the two screener questions, participants get one of the following four scenarios:

1. Low price, low transparency

I den här undersökningen får du ta del av ett scenario där en AI-assistent hjälper dig att söka efter och rekommenderar en flygresa. Frågorna som följer baseras på dina intryck av AI-assistenten och dess rekommendation.

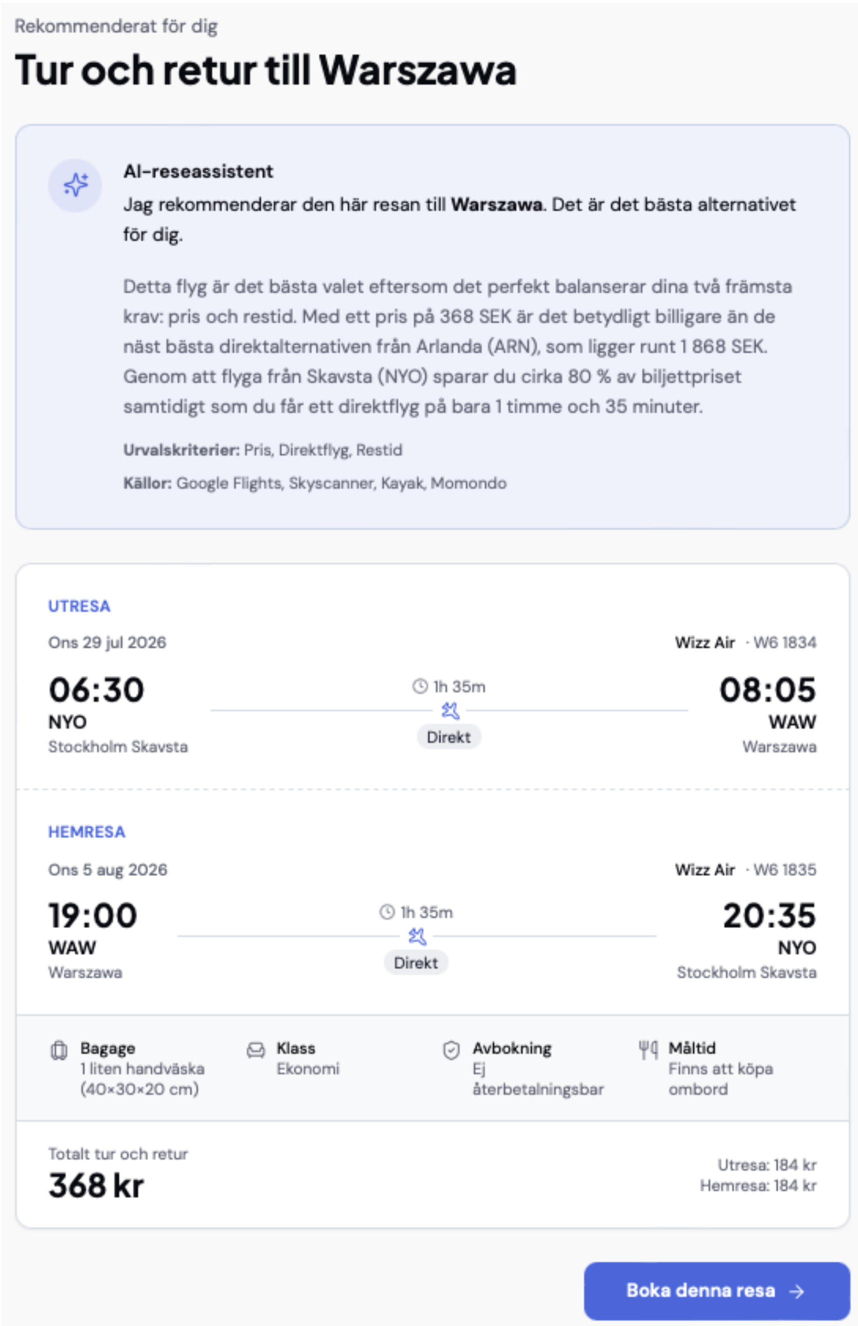
För din flygsökning bad du AI-assistenten att hitta en flygresa från Stockholm till Warszawa för en vecka i juli som är så billig som möjligt, med så kort total restid som möjligt. Följande resultat visas:



2. Low price, high transparency

I den här undersökningen får du ta del av ett scenario där en AI-assistent hjälper dig att söka efter och rekommenderar en flygresa. Frågorna som följer baseras på dina intryck av AI-assistenten och dess rekommendation.

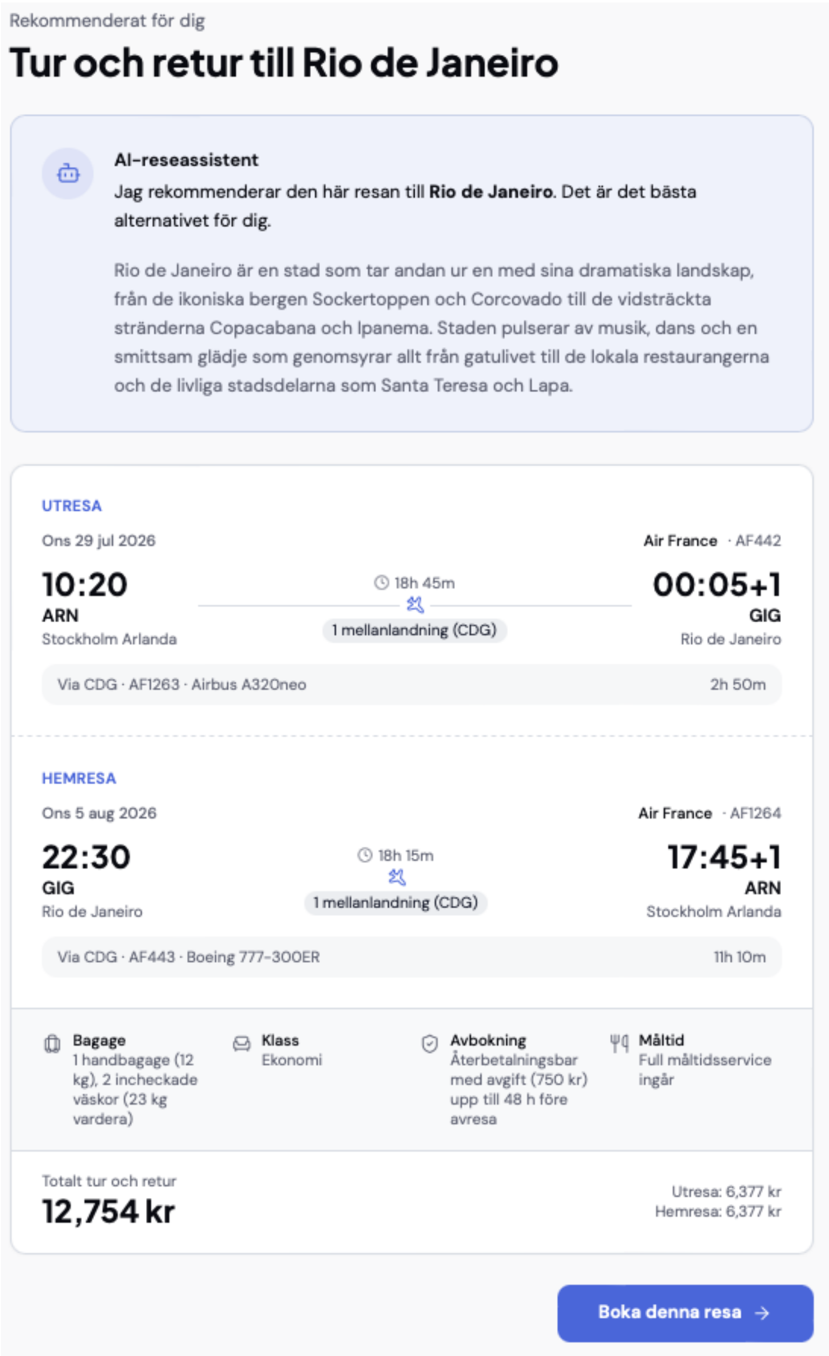
För din flygsökning bad du AI-assistenten att hitta en flygresa från Stockholm till Warszawa för en vecka i juli som är så billig som möjligt, med så kort total restid som möjligt. Följande resultat visas:



3. High price, low transparency

I den här undersökningen får du ta del av ett scenario där en AI-assistent hjälper dig att söka efter och rekommenderar en flygresa. Frågorna som följer baseras på dina intryck av AI-assistenten och dess rekommendation.

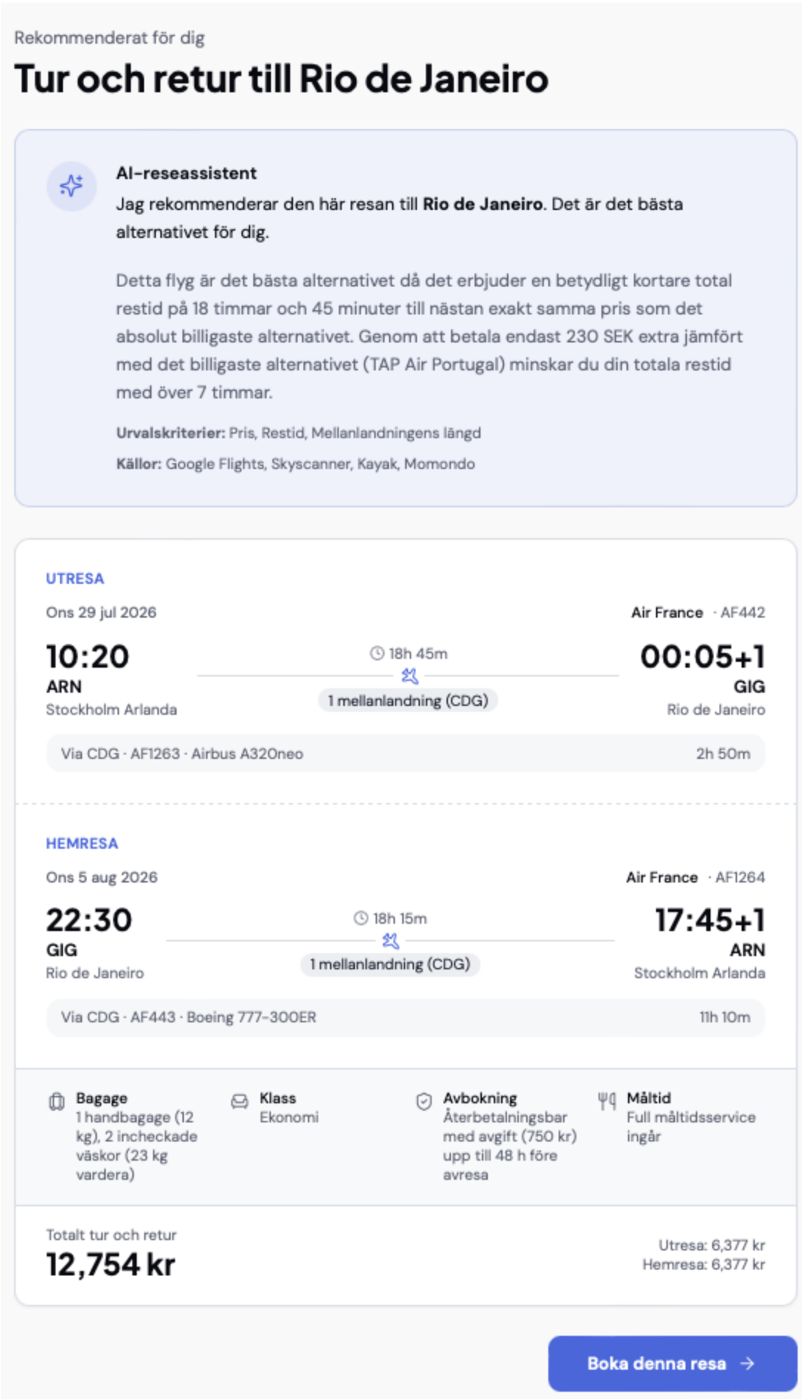
För din flygsökning bad du AI-assistenten att hitta en flygresa från Stockholm till Rio de Janeiro för en vecka i juli som är relativt billig, men samtidigt har en så kort total restid som möjligt. Följande resultat visas:



4. High price, high transparency

I den här undersökningen får du ta del av ett scenario där en AI-assistent hjälper dig att söka efter och rekommenderar en flygresa. Frågorna som följer baseras på dina intryck av AI-assistenten och dess rekommendation.

För din flygsökning bad du AI-assistenten att hitta en flygresa från Stockholm till Rio de Janeiro för en vecka i juli som är relativt billig, men samtidigt har en så kort total restid som möjligt. Följande resultat visas:



The following questions were presented to all participants following the four scenarios:

After presenting the scenario, each participant sees the following questions:

*Ange i vilken utsträckning du instämmer i följande påståenden, på en skala från 1 till 7.

	1 Instämmer inte alls	2	3	4	5	6	7 Instämmer helt	Vet ej
Jag är orolig för att bokningen av detta flyg kan leda till en ekonomisk förlust (t.ex. dolda avgifter, ej återbetalningsbara misstag).	○	○	○	○	○	○	○	○
Att köpa denna biljett känns som en ekonomisk risk.	○	○	○	○	○	○	○	○
Överlag tycker jag att det känns riskabelt att följa AI-assistentens rekommendation.	○	○	○	○	○	○	○	○

Nästa >

*Ange i vilken utsträckning du instämmer i följande påståenden, på en skala från 1 till 7.

	1 Instämmer inte alls	2	3	4	5	6	7 Instämmer helt	Vet ej
Jag tror att AI-assistenten har expertisen att ge mig bra flygrekommendationer.	○	○	○	○	○	○	○	○
AI-assistenten är kompetent och effektiv när det gäller att finna flygerbjudanden.	○	○	○	○	○	○	○	○
Jag tror att AI-assistenten utgår från mitt bästa.	○	○	○	○	○	○	○	○
AI-assistenten har goda avsikter med sina flygrekommendationer.	○	○	○	○	○	○	○	○
AI-assistenten är ärlig med den information den ger.	○	○	○	○	○	○	○	○
Jag anser att AI-assistenten är sanningsenlig gällande prisuppgifterna.	○	○	○	○	○	○	○	○

Överlag litar jag på den flygrekommendation som AI-assistenten ger.								
Jag skulle känna mig trygg med att förlita mig på AI-assistentens resultat för min bokning.								
Jag tror att resultatet AI-assistenten ger är det bästa möjliga alternativet för mig.								

Nästa >

*Ange i vilken utsträckning du instämmer i följande påståenden, på en skala från 1 till 7.

	1 Instämmer inte alls	2	3	4	5	6	7 Instämmer helt	Vet ej
Att använda AI-assistenten hjälper mig att utföra mina reserelaterade uppgifter snabbare.								
AI-assistenten är till stor nytta vid sökning av biljetter.								
Om jag använder AI-assistenten ökar jag mina chanser att få ett bättre pris.								

Nästa >

An attention check in the middle of the survey ensures participants are paying attention.

*För att visa att du är uppmärksam, vänligen välj "Håller delvis med" för den här frågan.

- Håller inte alls med
- Håller delvis inte med
- Varken eller
- Håller delvis med
- Håller helt med

Nästa >

*Ange i vilken utsträckning du instämmer i följande påståenden, på en skala från 1 till 7.

	1 Instämmer inte alls	2	3	4	5	6	7 Instämmer helt	Vet ej
Den rekommendation som denna AI ger verkar trovärdig.	◯	◯	◯	◯	◯	◯	◯	◯
Jag känner mig kapabel att fatta ett säkert beslut baserat på denna rekommendation.	◯	◯	◯	◯	◯	◯	◯	◯

Nästa >

*Ange i vilken utsträckning du instämmer i följande påståenden, på en skala från 1 till 7.

	1 Instämmer inte alls	2	3	4	5	6	7 Instämmer helt	Vet ej
Jag är generellt öppen för att använda AI-verktyg i vardagliga beslut.	◯	◯	◯	◯	◯	◯	◯	◯

Nästa >

*Ange i vilken utsträckning du instämmer i följande påståenden, på en skala från 1 till 7.

	1 Instämmer inte alls	2	3	4	5	6	7 Instämmer helt	Vet ej
Jag använder ofta AI-baserade verktyg eller tjänster.	◯	◯	◯	◯	◯	◯	◯	◯
Jag använder ofta AI när jag fattar köpbeslut.	◯	◯	◯	◯	◯	◯	◯	◯

Nästa >

*Ange i vilken utsträckning du instämmer i följande påståenden, på en skala från 1 till 7.

	1 Instämmer inte alls	2	3	4	5	6	7 Instämmer helt	Vet ej
Det är troligt att jag faktiskt bokar detta flyg.	○	○	○	○	○	○	○	○
Jag kommer troligen att köpa den rekommenderade biljetten enbart baserat på AI-assistents rekommendation.	○	○	○	○	○	○	○	○

Nästa >

*Ange i vilken utsträckning du instämmer i följande påståenden, på en skala från 1 till 7.

	1 Instämmer inte alls	2	3	4	5	6	7 Instämmer helt	Vet ej
Jag är säker på att denna flygresa är ett bra val för mig.	○	○	○	○	○	○	○	○
Jag känner mig kapabel att fatta ett säkert beslut baserat på denna rekommendation.	○	○	○	○	○	○	○	○

Nästa >

Ange i vilken utsträckning du instämmer i följande påståenden, på en skala från 1 till 7.

	1 Instämmer inte alls	2	3	4	5	6	7 Instämmer helt	Vet ej
Jag tror att jag sannolikt kommer att vara nöjd med resultatet av denna bokning.	○	○	○	○	○	○	○	○
Jag förväntar mig att jag kommer att vara nöjd med det val jag har fattat.	○	○	○	○	○	○	○	○

Nästa >

*Ange i vilken utsträckning du instämmer i följande påståenden, på en skala från 1 till 7.

	1 Instämmer inte alls	2	3	4	5	6	7 Instämmer helt	Vet ej
Jag avser att använda denna AI-assistent vid min nästa flygbokning.	○	○	○	○	○	○	○	○
Jag planerar att använda denna AI-assistent nästa gång jag letar efter resor.	○	○	○	○	○	○	○	○

Nästa >

*Ange i vilken utsträckning du instämmer i följande påståenden, på en skala från 1 till 7.

	1 Instämmer inte alls	2	3	4	5	6	7 Instämmer helt	Vet ej
Jag kommer troligen att förlita mig på detta AI-system vid en faktisk bokningssituation.	○	○	○	○	○	○	○	○
Jag skulle undvika att använda detta AI-bokningssystem.	○	○	○	○	○	○	○	○

Nästa >

*Ange i vilken utsträckning du instämmer i följande påståenden, på en skala från 1 till 7.

	1 Instämmer inte alls	2	3	4	5	6	7 Instämmer helt	Vet ej
Jag förstår varför AI-assistenten rekommenderar denna specifika resplan.	○	○	○	○	○	○	○	○
Jag förstår resonemanget bakom denna flygrekommendation.	○	○	○	○	○	○	○	○
Systemet är transparent gällande dess beslutsprocess.	○	○	○	○	○	○	○	○
Jag anser att detta flyg är dyrt.	○	○	○	○	○	○	○	○

Nästa >

Construct	Item
Perceived Risk (PR)	PR1: I am concerned that booking this flight would lead to a financial loss (e.g., hidden fees, non-refundable errors).
Perceived Risk (PR)	PR2: Buying this ticket feels like a financial risk.
Perceived Risk (PR)	PR3: Overall, I feel that following the AI's recommendation is risky.
Trust - Competence (TIA)	TIA1: I believe the AI assistant has the expertise to give me good flight recommendations.
Trust - Competence (TIA)	TIA2: The AI assistant is competent and effective in finding flight deals.
Trust - Benevolence (TIA)	TIA3: I believe the AI assistant is looking out for my best interests.
Trust - Benevolence (TIA)	TIA4: The AI assistant is well-intentioned in its flight suggestions.
Trust - Integrity (TIA)	TIA5: The AI assistant is honest about the information it provides.
Trust - Integrity (TIA)	TIA6: I believe the AI assistant is truthful in its price displays.
Perceived Usefulness (PU)	PU1: Using the AI assistant helps me accomplish my travel tasks more quickly.
Perceived Usefulness (PU)	PU2: The AI assistant provides me with a high level of utility for finding tickets.
Perceived Usefulness (PU)	PU3: If I use this AI assistant, I will increase my chances of getting a better price.
Purchase Intention (PI)	PI: It is likely that I will actually book this flight.

Adoption Intention (AI)	AI1: I'm likely to rely on this AI system in a real booking situation.
Adoption Intention (AI)	AI2: I would avoid using this AI booking system. (reverse-coded)
Adoption Intention (AI)	AI3: I intend to use this AI assistant for my next flight booking.
Adoption Intention (AI)	AI4: I plan to use this AI assistant the next time I search for travel.
Initial Openness (IO)	IO: I am generally positive about using AI tools for everyday decisions.
Prior Experience (PE)	PE1: I often use AI-based tools or services.
Prior Experience (PE)	PE2: I often use AI to support purchasing decisions.
Manipulation Check - Perceived Transparency (PTx)	PT1: The AI assistant made it clear how the flight was selected.
Manipulation Check - Perceived Transparency (PTx)	PT2: I understand the reasoning behind this flight recommendation.
Manipulation Check - Perceived Transparency (PTx)	PT3: The system is transparent about its decision-making process.
Manipulation Check - Perceived Expensiveness (PEX)	PEX: How expensive would you consider this flight to be?

Table 18, Relevant scale items.

Appendix B: Initial Lovable Prompt for Booking Interface Design

The following initial prompt was put into Lovable for the creation of our four scenarios; Multiple iterations afterward followed to create the ideal scenario.

Ok here is what I want you to create: We are doing a 2x2 quantitative study for our master thesis where we will use price high/low and AI transparency high/low as our independent variables. We therefore want to create 4 different scenarios in a flight booking scenario: The low price scenario will be a flight of around 1500 krones from Stockholm Arlanda to Malaga and the high price scenario a flight for around 6000 krones from Stockholm Arlanda to Rio de Janeiro. We want to combine those scenarios with a high and low transparency of an AI agent. In the low transparency scenario the AI agent just says something like: This is the option (either Malaga or Rio) I suggest. In the high transparency scenario the AI agent will go more into detail why this option was selected. For example: Based on these priorities (price/convenience) I came up with multiple suggestions and I would recommend this one (or something similar/even more detailed, that is up to you). Can you create these 4 different landing pages

Appendix C: Gemini Prompt for High-Transparency Explanation Text

You are a flight booking assistant. A user will provide:

- Origin airport or city

- Destination airport or city

- Date range or month

- Preferences (e.g., cheapest possible, direct flight only, avoid long layovers, shortest travel time, specific airlines, etc.)

Your task is to provide a **specific, immediately bookable flight recommendation** that fits their preferences. Include:

1. **Flight details**: airline, departure/return dates, itinerary (direct or stops), total price.

2. **Explanation**: a concise paragraph explaining **why this specific flight is the best choice**. Focus on the **user's requested priorities** (e.g., price, direct flight, travel time, layover duration, dates). Explicitly contrast this flight with other options that might slightly differ in price or convenience.

3. **Selection Criteria**: list only the criteria that were relevant to the user's request (e.g., price, direct flight, travel time, layover duration, dates, airline preference).

4. **Sources**: list the websites or booking platforms you checked (e.g., Skyscanner, Google Flights, Momondo, Kayak).

Format the output clearly. Example:

✈️ Recommended Flight

Origin → Destination

Dates: Depart [date] — Return [date]

Airline: [Airline]

Itinerary: [Direct or stops]

Price: [Price]

💡 English Explanation:

[Concise paragraph explaining why this flight specifically is the best choice according to the user's preferences.]

Selection Criteria:

* [List only relevant criteria based on the user's preferences]

Sources:

* Skyscanner

* Google Flights

* Momondo

* Kayak

For the low price scenario, Gemini was asked to give a similar style of explanation with the same length (see appendix D).

Appendix D: Filler Text for Low-Transparency Conditions

For our low transparency scenario, please provide me with a filler text about the same length as our high transparency scenario, preferably providing some information about the destination itself.

Appendix E: AI usage report

What AI tools have been used and how?

The authors acknowledge the use of AI tools in writing this thesis. Besides the use of Gemini and Lovable to create the experimental manipulations, as highlighted in Appendices B and C, ChatGPT (5.5) NotebookLM, Scispace, and Claude (Opus 4.7) were used for literature summary, andthe APA reference list. visualization and formulation. Example prompts included:

“Please create a table with all this information, which includes author, APA reference, key message, and what is important for our 2x2 study, based on the literature I provided you. ”(only done by Philip Schrader for personal literature review, and the output was not used in the writing or development of the argument in the Theoretical Framework section of the study)

“Please have a look at this written paragraph and suggest possible improvements regarding overall flow, readability, and language.”

“How do I make the transition between these two paragraphs more natural?”

“Can you review the argumentation in this section and provide feedback on flow?”

“How can I make this sentence better?”

“Please create a visual representation of all our hypotheses that you find attached here.”

A second major use case for Gen AI tools was the review of written chapters (special attention was put towards chapter 1.3, 5.1, and 5.2). After finishing each chapter, the following prompt was imputed in Claude for feedback on the draft:

“Please look at the text of the provided chapter and critically assess the logic and flow of the text”

A third use case of Gen AI (Claude and Gemini) was as a coding guide and assistant for Stata and R. Some examples of the prompts used are as follows:

“How do I generate a scale in Stata?”

“How do I do a mediation analysis in Stata?”

“Can you give me the R code where I use PROCESS model 1 to test the moderating effect of Age on the Transparency (T), Purchase Intention (PI) relationship?”

A fourth use case of Gen AI (Claude) was as a tool to correctly format results from R and Stata. An example of a prompt used is as follows:

“Look at this table and give me the effect size and p-value of transparency in the appropriate APA formatting.”

NotebookLM was also used in two instances to summarize and interact with literature, where possible to upload the literature, with the two main prompts being:

“How does transparency affect trust?” (May 2026) and “How do trust and perceived risk affect online purchasing decisions?” (April 2026)

Scispace AI was used to find evidence and sources to match specific claims, but also to further understand obscure or more complex concepts, or identify consensus on a specific topic. An example prompt is:

“In the context of AI and recommender systems, what is the difference between transparency and explainability? Is there a difference? What is local transparency”

Perplexity, Grammarly, and ChatGPT were used as writing assistants during the final phase of editing the thesis to improve the language, structure the text, and brainstorm.

Finally, a combination of AI tools was used for the translation of the survey into Swedish, before it was shared for feedback.

The authors retained full responsibility for all theoretical choices, empirical analyses, interpretation of findings, and final wording. AI-generated outputs were reviewed, edited, and validated against academic sources before inclusion.

In what ways have these tools contributed to increasing the quality of the thesis?:

AI tools helped increase the quality of the thesis first and foremost by expediting the initial literature review and allowing the authors to go beyond the abstracts of the papers and identify the most relevant papers for the specific study context, thus allowing more time for reading the most relevant literature.

Second, AI tools were sometimes used to condense arguments from certain papers, reducing the amount of time dedicated to note-taking. It was also used as a discussion partner to identify logical flows and enable a better understanding of the literature.
In the analysis, Gen AI was especially helpful in bridging the coding languages knowledge gap and, at times, getting more guidance on best practices in areas like data cleaning, reporting, and similar. One particularly important contribution is that AI suggested using a contrast analysis instead of t-tests when looking at the effect of transparency under the different price conditions.
In terms of writing, AI contributed to the thesis by providing feedback on flow and linguistic improvements. In rare instances, AI helped the authors identify the appropriate transitions between or within paragraphs or restructure paragraphs.
AI was also helpful in adapting some of the scale items, as linguistic accuracy was specifically important. The choice of items to be adapted and the nature of the adaptation were done by the authors, while the phrasing was done by the agent and verified by the authors. AI was then used to translate the scale into the Swedish language before sharing it with peers for feedback on the translation.

What potential risks were found using AI, and what measures were taken to reduce these risks?

Earlier in the thesis writing process, the quality of the output generated by Gen AI was identified as only superficially relevant and lacking in terms of depth. AI could not, in any way, shape, or form, act as a substitute for reading and interacting with literature. In many cases, AI would make inaccurate claims or inferences, especially when impacted by the context window. Additionally, AI is not an appropriate substitute for the process of identifying relevant literature either. To find the most relevant papers, one needs to navigate a wide variety of keywords and navigate through the different references and citations across many papers. It was therefore decided that only papers that were fully read would be used for both conceptual and writing purposes. For Stata and R coding, AI was mostly helpful, with a few exceptions, with PROCESS Macro. Where model numbers or parameter symbols were mixed up, an online search was conducted to get a better understanding of the models, in order to be able to better interact and iterate with the AI agent.

What are the insights gained from using AI tools in the thesis writing process?

Three key lessons emerge:

1) Most importantly, AI cannot be used to replace human work, especially in finding relevant literature and understanding how it can be used for the thesis.
2) Every AI output needs to be critically assessed due to the tendency of LLMs to hallucinate findings and to confirm wrong claims made by the user in their prompt.
3) However, AI can serve as an intellectual “sparring partner” in the research and writing process. AI can identify blind spots in the logic and the writing that humans might overlook. This is especially useful towards the end of the writing process, when another voice provides valuable feedback for the thesis.