

Varying transparency in AI chatbots: Implications for perceived trust, perceived expertise, and behavioral intention.

Varying the transparency levels of AI chatbots through explainability and accountability. Observing implications for behavioral intention and the possible moderating effect of perceived trust and perceived expertise.

Student Name: Tasio LOPEZ--HEURTIN

Student Number: 535928

Supervisor: Dr. Niels Vink

Master Media Studies - Media & Business

Erasmus School of History, Culture and Communication

Erasmus University Rotterdam

Master's Thesis

June 2024

Word Count: 12187

Varying transparency in AI chatbots: Implications for perceived trust, perceived expertise, and behavioral intention.

ABSTRACT

The rapid expansion of Artificial Intelligence (AI) has significantly transformed various sectors, including banking, where AI chatbots are revolutionizing service delivery. Despite their increasing adoption, comprehensive research on how transparency influences perceived trust and expertise, and how these factors affect users' behavioral intentions to use financial services, remains limited. This study aims to fill this gap by examining these relationships in the context of AI banking chatbots.

Transparency in AI systems involves clearly communicating the AI's capabilities, limitations, and decision-making processes. A key aspect of transparency is explainability, which refers to the AI's ability to articulate its actions and decisions in a way that users can understand. Accountability is another crucial component, ensuring that AI systems are responsible for their actions by providing mechanisms for feedback, issue reporting, and redress. Perceived trust is defined as the user's belief in the AI's reliability and beneficial intent, while perceived expertise reflects confidence in the chatbot's competence. Behavioral intention refers to the likelihood that a user will engage with services suggested by the chatbot.

This study employed a quantitative approach, gathering data from 273 participants through a survey that measured explainability, accountability, perceived trust, perceived expertise, and behavioral intention. Data analysis included linear regression analyses, ANOVAs, moderation analyses and mediation analyses.

Key findings revealed that perceived expertise was a strong predictor of behavioral intention, while perceived trust also significantly influenced behavioral intention. However, the mediation analysis indicated that perceived trust does not mediate the relationship between perceived expertise and behavioral intention.

Interestingly, the moderation analysis showed that perceived trust does not moderate the relationship between explainability / accountability and behavioral intention. Instead, perceived expertise directly influenced behavioral intention, independent of perceived trust. These findings underscore the critical role of perceived expertise in driving user adoption of AI chatbots in the financial sector.

The study contributes to the existing literature by highlighting the importance of perceived expertise and trust in influencing behavioral intention, while challenging the expected mediation role of trust. For practical applications, these insights can guide banking institutions in designing and implementing AI chatbots that enhance user engagement through demonstrating high expertise and building trust.

KEYWORDS: *Explainability, Accountability, Behavioral intention, Perceived Trust, Perceived Expertise*

Table of contents

ABSTRACT.....	2
Introduction.....	6
Conceptual model	7
Structure of the paper	9
Theoretical Framework.....	10
Transparency.....	10
Explainability.....	11
Accountability.....	13
Behavioral intention to use financial services	14
Perceived trust.....	15
Perceived expertise	17
Combined effect of trust and expertise	17
Trust as a mediator between expertise and behavioral intention	18
Methods.....	20
Research Design.....	20
Participants.....	20
Materials	24
Perceived trust.....	25
Perceived expertise	26
Behavioral intention to use of financial services	26
Overview of scales	26
Data quality control.....	27
Factor Analysis:	27
Perceived Trust	27
GEM / General Expertise Measurement	28
Behavioral intention to use financial services	29
Procedure	30

Ethical Considerations	31
Results:.....	33
Overview of hypotheses:.....	42
Conclusion	43
Overview of research findings	43
Hypothesis 1: Transparency of AI chatbots and behavioral intention.....	43
Hypothesis 2: Perceived trust as a moderator	43
Hypothesis 3: Perceived expertise as a moderator.....	44
Hypothesis 4: Combined effect of perceived trust and expertise	44
Hypothesis 5: Mediation effect of perceived trust.....	44
Discussion	45
Academical and professional implications	46
Academically	46
Professionally.....	48
Limitations and future research	49
Conclusion	50
Literature.....	51
Appendix A: Questionnaire	56
Appendix B: Stimuli / Fictional conversations.....	75
1. Low Explainability and low accountability	75
2. Low explainability and high accountability.....	79
3. High explainability and low accountability	83
4. High explainability and high accountability	87
Appendix C: ChatGPT prompts for fictional conversations.....	91

Introduction

The recent expansion of Artificial Intelligence (AI) has brought significant transformations across various sectors, with the banking industry being no exception (Fares et al., 2023, p.835). AI chatbots have emerged as a promising technology, revolutionizing the way banking services are delivered (Mori & Du, 2023, p.3). These chatbots provide round-the-clock service, handle a multitude of customer queries, and perform transactions with efficiency and precision. Despite the growing adoption of AI chatbots in banking, there is still a lack of comprehensive research exploring the influence of transparency on perceived trust and perceived expertise, and how these factors affect a user's behavioral intention to use financial services (Choung et al., 2022, p.9).

Transparency in AI systems is a multifaceted concept (Walmsley, 2021, p.590) that involves the clear communication of the AI's capabilities, limitations, and decision-making processes to the user (Liu, 2021, p.385). This transparency is crucial as it can significantly influence the user's trust in the system and their willingness to engage with it. In the context of AI chatbots, transparency refers to how well users understand the functioning and decision-making processes of the chatbot, which includes the data being used by the AI and the reasoning behind the AI's responses.

Explainability is a critical component of transparency. It refers to the AI system's ability to articulate its actions and decisions in a manner that is comprehensible to the user (Abu-Rasheed et al., 2024, p.6). Explainability helps demystify the AI's operations, making users more comfortable and confident in interacting with the chatbot. For instance, in the context of banking, a chatbot should not only provide a recommendation for a financial product but also explain why that product is suitable for the user based on their financial history and goals.

Accountability is another essential element of transparency. It involves the AI system being responsible for its actions and decisions, ensuring that users have mechanisms to provide feedback, report issues, and seek redress if the system makes a mistake (Novelli et al., 2023, p.3). In the banking sector, accountability can be demonstrated through clear communication about data usage, error rectification processes, and channels for user complaints and suggestions.

Perceived trust is the belief that the AI system will behave in a manner that is expected, reliable, and beneficial to the user (Lukyanenko et al., 2022, p.2000). Trust is a

pivotal factor in the acceptance and continued use of AI technologies. On the other hand, perceived expertise is defined as the user's belief in a chatbot's knowledge and competence (Kuhail et al., 202, p.9753). It reflects the user's confidence in the chatbot's ability to handle banking-related tasks effectively. Lastly, behavioral intention refers to the likelihood that a user will engage in a particular behavior (Gatzioufa & Saprikis, 2022, p.3), which in this case would be using financial services suggested by the chatbot.

This study aims to address this current gap in academic literature by exploring the relationship between these concepts in the context of AI banking chatbots. Existing studies acknowledge that while their findings indicate that lack of transparency can hinder perceived trust, there is still a scarcity of research connecting perceived trust to transparency and subsequent acceptance of AI systems (Choung et al., 2022, p.9). This highlights the need for more studies investigating the relationship between perceived trust, transparency, and behavior. This study will contribute to the current body of literature by adding the variable of perceived expertise in its moderation analysis and focusing on studying the phenomenon in the context of the banking industry.

Additionally, this research could have significant societal relevance for banking institutions. While most banks provide online services (Fares et al., 2023, p.836), they can directly use the insights about the effect of transparency on behavioral intention to tailor their chatbots and AI services accordingly. Understanding and fostering trust towards AI chatbots, particularly through enhanced transparency, can be crucial for user acceptance and play an important role in guiding behavioral intention (Lukyanenko et al., 2022, p.2006).

Therefore, the research question of this study is: *“How does transparency of AI chatbots affect a user's behavioral intention to use financial services through perceived trust and perceived expertise?”*. By exploring this question, the study aims to provide valuable insights into the dynamics of user interaction with AI chatbots in the banking sector. Furthermore, this will lead to practical recommendations for improving user experience and enhancing the effectiveness of AI applications in financial services.

Conceptual model

The conceptual model guiding this study integrates the constructs of transparency, explainability, accountability, perceived trust, perceived expertise, and behavioral intention to explore their interrelationships within the context of AI chatbots in banking (see Figure 1.1).

Transparency is the overarching concept, encompassing explainability and accountability, which are varied in this study. Explainability refers to the AI's ability to provide understandable and clear explanations for its decisions and actions, thereby helping users comprehend the reasoning behind the chatbot's responses. Accountability, on the other hand, involves the chatbot's responsibility for its actions, including mechanisms for feedback, error correction, and user redress.

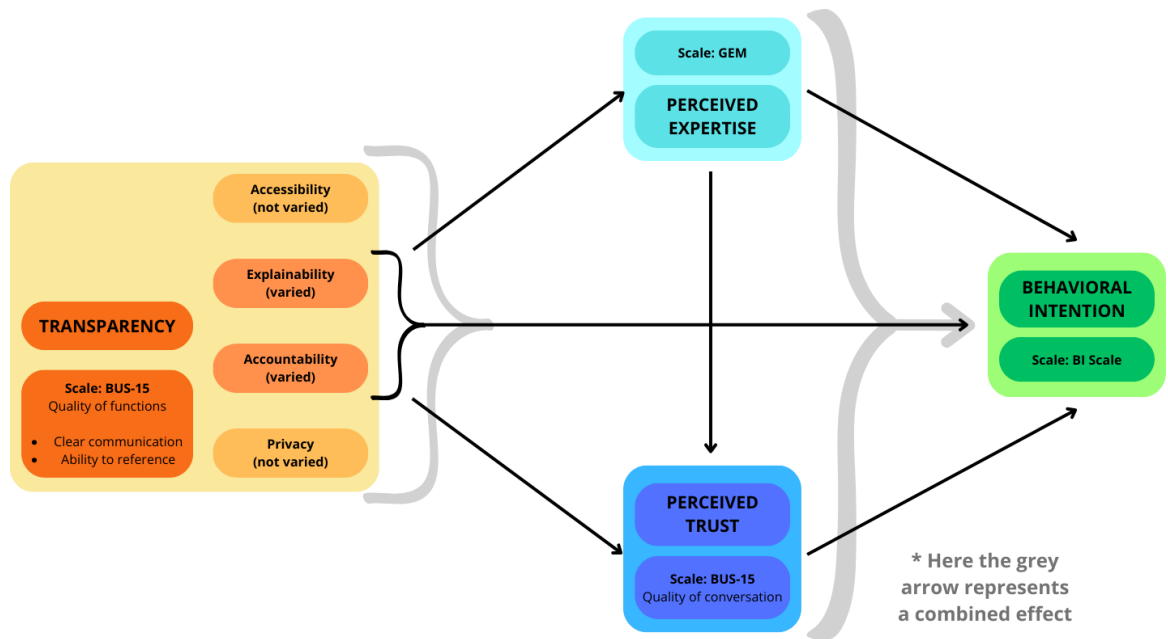


Figure 1.1: Conceptual model

Perceived trust and perceived expertise are ~~posited as~~ critical mediators in this model. Perceived trust is the user's belief in the reliability and beneficial intent of the chatbot, which can be enhanced through transparent practices. Perceived expertise reflects the user's belief in the chatbot's knowledge and competence, which can be influenced by how well the chatbot explains its decisions and demonstrates accountability.

The model suggests that higher levels of transparency, achieved through improved explainability and accountability, will enhance perceived trust and perceived expertise. These, in turn, are expected to positively influence the user's behavioral intention to engage with the chatbot for financial services. The combined effect of perceived trust and perceived expertise is also hypothesized to have a stronger impact on behavioral intention than either factor alone. Finally, perceived trust is expected to mediate the relationship between perceived expertise and behavioral intention.

Structure of the paper

The remainder of this thesis will first present an in-depth review of existing literature regarding the concepts central to this study, and the theoretical underpinnings and interrelationships of these concepts. After this, we will provide details about the methods employed for the collection, preparing, and analysis of the data. This section also discusses the survey construction and how concepts were manipulated in the stimuli. After this, the results of the data analysis will be presented before discussing the implications of these results. Finally, we will provide a comprehensive interpretation of the findings, before discussing the limitations of the study, and provide recommendations for future research.

Theoretical Framework

Transparency

Transparency, in the context of Artificial Intelligence (AI) systems, is a multifaceted concept that has gained significant attention in recent years due to the rapid advancement and widespread adoption of AI technologies (Liao & Vaughan, 2024, p.2; Walmsley, 2021, p.586). It is particularly relevant in the realm of AI chatbots, where the clarity and understandability of the system's operations and decision-making processes can significantly influence user trust and behavioral intention (Walmsley, 2021, p.586).

At its core, transparency in AI systems refers to the extent to which the system's operations, decision-making processes, and underlying algorithms are made clear and understandable to its users (Balasubramanian et al., 2023, p.6). This involves not only the disclosure of technical details but also the communication of the system's capabilities, limitations, and the reasoning behind its actions.

Several key components contribute to the overall transparency of an AI system:

- **Clear communication of capabilities:** The AI system should clearly communicate its capabilities and limitations to the user. This includes explaining what the system can do, what data it uses, and how it makes decisions. This aspect of transparency is crucial as it sets the user's expectations and helps them understand the system's strengths and limitations (Balasubramanian et al., 2023, p.6).
- **Explainability:** The AI system should be able to explain its actions and decisions in a way that is understandable to the user. This involves providing clear, concise, and understandable explanations for its recommendations or actions. Explainability is a critical aspect of transparency as it helps users understand why the system behaves the way it does and fosters trust in the system (Balasubramanian et al., 2023, p.6).
- **Data privacy and security:** The AI system should clearly communicate how it handles and protects user data. This includes explaining what data is collected, how it is used, who has access to it, and how it is protected. This aspect of transparency is particularly important in the context of AI chatbots, where sensitive user data is often involved (Balasubramanian et al., 2023, p.6).
- **Accountability:** The AI system should be accountable for its actions and decisions. This means that there should be mechanisms in place for users to provide feedback,

report issues, and seek redress if the system makes a mistake. Accountability is a crucial aspect of transparency as it ensures that the system is responsible for its actions and decisions (Balasubramanian et al., 2023, p.6).

The impact of transparency on user behavior is significant. A transparent AI system can foster trust, which in turn can influence a user's behavioral intention to use the system (Buijsman, 2024, p.3). For instance, a study by Wanner et al. (2022, p.2095) found that higher transparency can lead to higher trust, which in turn leads to higher behavioral intention. This highlights the importance of transparency in AI systems and its role in influencing user behavior.

In the following sections, we will delve deeper into the concepts of explainability and accountability, two integral components of transparency, and discuss their roles in fostering user trust and influencing behavioral intention.

Explainability

Explainability, in the context of AI systems, is a critical aspect of transparency that refers to the ability of an AI system to articulate its actions and decisions in a manner that is comprehensible to the user (Balasubramanian et al., 2023, p.6). It is a fundamental requirement for fostering user trust and satisfaction (Chazette, 2021, p.6), which are key indicators of chatbot quality.

The importance of explainability in AI systems has gained significant attention (Miller, 2019, p.1) and has been emphasized in recent academic literature (Shin, 2020, p.542); Shin, 2021, p.1048). For instance, a systematic review by Kuhail et al. (2023, p.975) analyzes 36 papers on educational chatbots to understand their design principles, interaction styles, and empirical evidence. The study found that chatbots which provide clear and understandable explanations significantly improve learning outcomes and student satisfaction. Specifically, it highlights how personalized interactions and immediate feedback from chatbots contribute to a better learning experience.

Another study by Abu-Rasheed et al. (2024, p.6) explores the potential of chatbots to engage students in conversations, similar to discussions with peers or mentors, which is crucial for explainability. The researchers designed a chatbot module using large language models (LLMs) and knowledge graphs to provide conversational explainability. Through a

user study, they found that chatbots capable of delivering clear explanations effectively support student decisions on learning recommendations, enhancing their engagement and learning outcomes. The study involves using LLMs for generating relevant explanations and connecting students with human mentors for additional support (Abu-Rasheed et al., 2023, p.3). This study underscores the role of explainability in fostering student engagement and improving decision-making in learning contexts.

Explainability is particularly relevant in the context of banking chatbots, where users often need to understand complex financial information and make important financial decisions. A chatbot that can clearly communicate its actions and decisions and provide easy-to-understand responses is likely to be perceived as more transparent and of higher quality by users. Transparency, including its facets of explainability in this case, is crucial for both practical functionality and business success. Practical guidelines for implementing explainable AI, as discussed by Balasubramaniam et al. (2023, p.6), Grennan et al. (2022, p.2), highlight the importance of clear and understandable AI systems in achieving business goals and fostering user trust.

However, achieving high explainability in a banking chatbot involves several key factors (Balasubramaniam et al. 2023, p.8, Miller, 2019, p.27):

- **Clarity of explanations:** The chatbot should provide clear and concise explanations for its actions and decisions. This includes explaining the reasoning behind its recommendations or actions in a way that is understandable to the user.
- **Contextual relevance:** The chatbot's explanations should be relevant to the user's context. This means that the chatbot should consider the user's financial situation, needs, and preferences when providing explanations.
- **Consistency:** The chatbot's explanations should be consistent across different interactions. This means that the chatbot should provide similar explanations for similar actions or decisions, which can help users understand the chatbot's behavior over time.

Overall, explainability is a critical aspect of transparency in AI systems that can significantly influence user trust and behavioral intention to use financial services (Miller, 2019, p.1). This can suggest that future research should continue to explore the role of explainability in AI systems, and develop strategies to enhance explainability in banking chatbots.

Accountability

Accountability, in the context of AI systems, is a critical aspect of transparency that refers to the responsibility of an AI system for its actions and decisions (Balasubramaniam et al, 2022, p.6) It includes mechanisms for users to provide feedback, report issues, and seek redress if the system makes a mistake. This aspect of transparency is integral to the perceived quality of a chatbot.

The importance of accountability in AI systems has been emphasized in recent academic literature. For instance, a study by Novelli, Taddeo, and Floridi (2023, p.3) discusses accountability in artificial intelligence and how it works. They argue that accountability in AI is often defined too imprecisely because its multifaceted nature and the sociotechnical structure of AI systems imply a variety of values, practices, and measures to which accountability in AI can refer. They address this lack of clarity by defining accountability in terms of answerability, identifying three conditions of possibility: authority recognition, interrogation, and limitation of power (Novelli et al., 2023, p.3). Specifically, they outline an architecture of seven features: context, range, agent, forum, standards, process, and implications (Novelli et al., 2023, pp. 4-5).

To create more accountable AI chatbots, these features can be operationalized by clearly defining the chatbot's roles and responsibilities, implementing mechanisms for users to question and receive explanations about the chatbot's decisions, establishing boundaries for decision-making capabilities, and tailoring the chatbot's functionalities to specific contexts such as education or healthcare. Additionally, assigning accountability to developers and operators, establishing feedback and redress forums, adhering to legal and ethical standards, and defining clear consequences for the chatbot's actions are crucial steps. By incorporating these elements, AI chatbots can enhance user trust and ensure ethical and responsible AI use (Abdul et al, 2018, p. 8).

Accountability is particularly relevant in the context of banking chatbots, where users often need to understand complex financial information and make important financial decisions. A chatbot that can effectively manage these aspects is likely to be perceived as more accountable and of higher quality by users (Diakopoulos, 2017, p.816).

Achieving high accountability in a banking chatbot thus involves several key factors (Novelli et al., 2023, p. 3; Diakopoulos, 2015, p.402):

- Feedback mechanisms: The chatbot should provide mechanisms for users to provide feedback and report issues. This includes providing options for direct feedback and ensuring that user concerns are addressed promptly.
- Redress mechanisms: The chatbot should provide mechanisms for users to seek redress if the system makes a mistake. This includes providing options for dispute resolution and ensuring that user concerns are addressed promptly.
- Responsibility for actions and decisions: The chatbot should be responsible for its actions and decisions. This means that the chatbot should take immediate and effective action to resolve issues and ensure that its actions and decisions are in line with the user's expectations.

In definitive, accountability is a critical aspect of transparency in AI systems that can significantly influence user trust and behavioral intention to use financial services (Lepri et al., 2018, p. 617). Abdul et al. (2018, p. 16) further highlight the importance of involving multidisciplinary teams in the development and evaluation of accountable AI systems, ensuring diverse perspectives and robust accountability mechanisms.

Behavioral intention to use financial services

Behavioral intention, particularly in the context of AI chatbots, refers to the likelihood that a user will engage in a particular behavior, in this case, the use of financial services through AI chatbots. It is a key determinant of user acceptance and use of AI technologies.

The concept of behavioral intention in AI systems has been extensively studied in recent academic literature. For instance, a study by Gatzoufa and Saprikis (2022) conducted a comprehensive literature review on users' behavioral intention toward chatbots' adoption. Their findings identified distinct categorization criteria, including research field, applied theoretical models, research types, methods, and statistical measures, as well as factors affecting the intention to adopt and use chatbots. (Gatzoufa & Saprikis, 2022, p. 3). They highlighted the importance of performance expectancy, effort expectancy, and social influence as significant predictors of users' behavioral intentions (Gatzoufa & Saprikis, 2022, p. 4).

Another study by Wu et al. (2022) explored the factors influencing the willingness to accept AI-assisted learning environments. Using a quantitative research design, they surveyed

500 college students to examine the relationships between effort expectancy, performance expectancy, social influence, and the students' willingness to accept AI-assisted learning environments. The results indicated that all three factors were positively related to willingness. Specifically, effort expectancy, which refers to the ease of use associated with the AI-assisted environment, significantly influenced students' acceptance (Wu et al., 2022, p.15). Performance expectancy, defined as the perceived benefits and improvements in learning outcomes from using AI, also showed a strong positive correlation with willingness (Wu et al., 2022, p.16). Additionally, social influence, or the degree to which students perceive that important others believe they should use the AI-assisted learning environment, was found to be a significant predictor of acceptance (Wu et al., 2022, p.17). These findings underscore the importance of designing AI-assisted learning tools that are user-friendly, demonstrably beneficial, and socially endorsed to enhance acceptance among college students.

In the context of banking chatbots, behavioral intention is particularly important. Users are likely to use a chatbot for financial services if they perceive it as transparent, trustworthy, and competent. Moreover, a user's behavioral intention can be influenced by their perception of the chatbot's expertise and trustworthiness.

Based on these insights, the following hypothesis was developed:

H1: Higher transparency of AI chatbots leads to a higher intention to use financial services.

Perceived trust

Perceived trust, particularly in the context of AI chatbots, is a multifaceted construct that involves the belief that AI systems will behave in a manner that is expected, reliable, and beneficial to the user. It is a key determinant of user acceptance and use of AI technologies.

The concept of trust in AI systems has been extensively studied in recent academic literature. For instance, a study by Yang and Wibowo (2022) developed a comprehensive conceptual framework to understand users' trust in AI systems. This framework was based on a systematic review of 131 studies conducted between 2015 and 2022, identifying key components and influencing factors of user trust in AI (Yang & Wibowo, 2022, p. 2054). They identified several critical factors influencing trust, including the transparency of the

system, the perceived expertise of the AI, and the user's prior experiences and predispositions. Transparency was highlighted as a major factor, emphasizing that clear, understandable explanations of AI operations significantly enhance user trust (Yang & Wibowo, 2022, p. 2060). The perceived expertise of the AI, referring to its ability to perform tasks accurately and reliably, was also found to be crucial (Yang & Wibowo, 2022, p. 2061). Additionally, users' prior experiences and predispositions towards technology play a significant role in shaping their trust in AI systems (Yang & Wibowo, 2022, p. 2062). The study concluded that addressing these factors through improved design and communication strategies can significantly enhance user trust in AI.

In the context of banking chatbots, perceived trust is particularly important. Users are likely to trust a chatbot if they perceive it as transparent and understand its functioning and decision-making process. Moreover, a user's trust in a chatbot can be influenced by their perception of the chatbot's expertise. If users perceive the chatbot as knowledgeable and competent in handling banking-related tasks, they are more likely to trust it.

Another study by Wanner et al. (2022) highlighted the role of trust as a moderator in the relationship between transparency and behavioral intention. The study operationalized 'behavioral intention' by measuring users' likelihood to continue using, recommending, and relying on the AI system in the future. Using a mixed-methods approach, they surveyed 400 participants to assess their perceptions of AI transparency, trust, and behavioral intention (Wanner et al., 2022, p. 2080). They found that higher transparency can lead to higher trust, which in turn leads to higher behavioral intention. Specifically, transparency was shown to enhance users' trust in the system by providing clear and understandable explanations of its operations, which subsequently increased their intention to use and recommend the system (Wanner et al., 2022, pp. 2095-2096). This suggests that trust plays a crucial role in mediating the relationship between transparency and behavioral intention, underscoring the importance of designing transparent AI systems to foster user trust and engagement.

Based on these insights, the following moderation hypothesis was developed:

H2: Perceived trust moderates the relationship between transparency of AI chatbots and intention to use financial services.

Perceived expertise

Perceived expertise, particularly in the context of AI chatbots, refers to the user's belief in a chatbot's knowledge and competence (Yang & Wibowo, 2022, p. 2062). It is a critical factor in the acceptance and use of AI chatbots.

The concept of perceived expertise in AI systems has been extensively studied in recent academic literature. For instance, a study by Zhang et al. (2022, p.9) investigates how people trust and rely on an AI assistant that performs with different levels of expertise relative to the person, ranging from completely overlapping expertise to perfectly complementary expertise. They found that participants successfully perceived when the assistant was an expert or not within the same task and calibrate their reliance on the AI to improve team performance.

In the context of banking chatbots, perceived expertise is particularly important. As discussed by Yang & Wibowo (2022, p.2062), users are likely to trust a chatbot if they perceive it as knowledgeable and competent in handling banking-related tasks. Moreover, a user's trust in a chatbot can be influenced by their perception of the chatbot's expertise. If users perceive the chatbot as knowledgeable and competent in handling banking-related tasks, they are more likely to trust it. Similarly to perceived trust, this suggests that perceived expertise plays a crucial role in the inter-relationships between transparency, perceived trust and behavioral intention.

Based on these insights, the following moderation hypothesis was developed:

H3: Perceived expertise moderates the relationship between transparency of AI chatbots and intention to use financial services.

Combined effect of trust and expertise

The combined effects of perceived trust and perceived expertise on behavioral intention in AI systems is a complex interplay that can significantly influence user acceptance and use of AI technologies. This interaction effect suggests that the combined influence of trust and expertise on behavioral intention is greater than their individual effects alone.

In the context of banking chatbots, the interaction effects of perceived trust and perceived expertise on behavioral intention can potentially be important. Users are likely to

use a chatbot for financial services if they perceive it as trustworthy and competent. Moreover, the combined effect of trust and expertise can have a greater influence on behavioral intention than their individual effects.

Wanner et al. (2022) emphasize that trust moderates the relationship between transparency and behavioral intention. Their findings indicate that higher transparency enhances trust, which subsequently increases behavioral intention, highlighting the critical role of trust in mediating these relationships (Wanner et al., 2022, pp. 2095-2096). This enhancement of trust is crucial for its combined effect with perceived expertise on behavioral intention as pointed out by Yang & Wibowo (2022, p. 2060).

Additionally, Gatzidoufa and Saprikis (2022) identified factors, when combined with trust and expertise, can markedly enhance user behavioral intention (Gatzidoufa & Saprikis, 2022, pp. 4-5). This echoes with Wu et al. (2022) finding that effort expectancy, performance expectancy, and social influence positively correlate with users' willingness to accept AI-assisted environments. This underscores the importance of perceived expertise (performance expectancy) alongside trust in influencing behavioral intention (Wu et al., 2022, pp. 15-17).

Based on these insights, the following interaction hypothesis was developed:

H4: Behavioral intention to use financial services is highest, when both perceived trust and perceived expertise are high.

Trust as a mediator between expertise and behavioral intention

Additionally, another hypothesis was developed to add depth to the analysis of trust in the context of this study and its behavior when interacting with other main concepts studied and closely related according to literature.

Perceived expertise, which refers to users' belief in an AI chatbot's competence, is a crucial factor in fostering initial trust. According to the technology acceptance model (TAM), users are more likely to adopt technology they perceive as competent (Davis, 1989, p. 320). On the other hand, trust, a key determinant in technology adoption, involves beliefs in the technology's reliability, integrity, and competence (McKnight et al., 2002, p. 336). Trust reduces perceived risks and increases user willingness to engage with the technology (Pavlou & Fygenson, 2006, p. 118). Mayer, Davis, and Schoorman (1995, p. 712) suggest that trust is

built on perceptions of ability, benevolence, and integrity. In the context of AI chatbots, perceived expertise directly contributes to perceived ability, fostering trust.

The mediating role of trust can be expected because it acts as a bridge translating perceived expertise into behavioral intention. When users trust a chatbot's expertise, their concerns about accuracy and reliability diminish, enhancing their intention to use the chatbot (Gefen et al., 2003, p. 55; Yang & Wibowo, 2022, p. 2060). Previous studies have demonstrated that trust enhances user engagement and willingness to adopt AI systems, supporting the mediation hypothesis (Wanner et al., 2022, p. 2095).

Following these insights, the following hypothesis was drafted:

H5: Perceived trust mediates the relationship between perceived expertise and behavioral intention to use financial services.

Ultimately, this theoretical framework has explored the key concepts of transparency, explainability, accountability, perceived trust, perceived expertise, and behavioral intention in the context of AI chatbots in the banking industry. Each of these concepts could play a crucial role in influencing user acceptance and use of AI chatbots for financial services.

Methods

Research Design

A quantitative research design was employed for this study because of its ability to determine statistical relationships among variables and predict phenomena accurately. The combination of a quantitative approach, moderation and mediation analyses, and a scenario-based questionnaire offers a comprehensive and methodologically sound framework for addressing the research question (Lukyanenko et al., 2022, p. 1996; Choung et al., 2022, p. 4, Jafarkimi et al., 2016, p.148). By employing these methods, we ensure that the study is not only scientifically rigorous but also relevant and impactful, providing valuable insights into the design and implementation of transparent AI chatbots in the banking industry.

Furthermore, these approaches allow for precise measurement of variables and statistical analysis, which is essential for drawing valid conclusions. This methodological rigor is vital for studies aiming to understand complex interactions between variables (Creswell & Creswell, 2018, p.199). This approach helps in identifying causal relationships and understanding how different factors interact within a study (Field, 2018, p.58).

Participants

Participants for this study were recruited through paid survey platforms, ensuring the targeted sample met the study criteria. These platforms filter respondents based on user profiles and pre-existing information, safeguarding participants' anonymity while enabling efficient and effective survey distribution. Utilizing such services also benefits from the platform's built-in quality and attention checks, enhancing the validity and reliability of data collection.

The inclusion criteria for participants required them to be 18 years or older and possess a proficient level of English to accurately comprehend and respond to the survey questions. Familiarity with chatbots was assessed through three specific questions to ensure a diverse range of experiences: participants were asked how familiar they were with chatbots and conversational interfaces (on a 1-5 scale), whether they had used a chatbot or conversational interface before (on a 1-5 scale), and how often they used chatbots weekly (on

a 1-5 scale). Responses from test runs labeled with "test" and participants under 18 years old were excluded from the dataset to maintain data integrity.

The final sample consisted of 273 participants, with a majority identifying as female (61.9%). The most common educational status among participants was a bachelor’s degree (35.5%). Regarding familiarity with chatbots and conversational interfaces, the largest group reported being quite familiar (48.7%). Regarding familiarity with chatbots, nearly half of the respondents (48.7%) are quite familiar with them, while 28.2% have limited experience. A smaller segment is extremely familiar (14.3%), and 8.8% are not familiar at all. Usage of chatbots is relatively common, with 41.4% of respondents having definitely used them before and 39.6% using them sometimes. (see table 3.1).

Table 3.1

Socio-demographic characteristics

Variables	Characteristics	Frequency	Percentage
Gender	Male	104	38.1
	Female	169	61.9
Educational status	Primary school	1	0.4
	Secondary school / high school	53	19.4
	Vocational degree after high school	51	18.7
	Bachelor’s degree	97	35.5
	Master’s degree	55	20.1
	PhD, MBA, or other equivalent	10	3.7
	Other, namely	5	1.8
	Prefer not to say	1	0.4
Familiarity with chatbots and/ or other conversational interfaces	Limited experience	77	28.2
	Quite familiar	133	48.7
	Extremely familiar	39	14.3

	Not familiar at all	24	8.8
Use of a chatbot or a conversational interface before	Definitely no	23	8.4
	Almost never	25	9.2
	I don't know	4	1.5
	Sometimes	108	39.6
	Definitely yes	113	41.4
How often do you use chatbots weekly?	Never	74	27.1
	Around once a week	123	45.1
	I don't know	40	14.7
	Every other day	26	9.5
	Daily	10	3.7

Overall, the mean age of the participants was 42.66 years ($SD = 18.51$). The median age was 38 years [IQR = 24, 60], with the ages ranging from 19 to 83 years, yielding a range of 64 years. An overview of the participant's age is provided in figure 3.1.

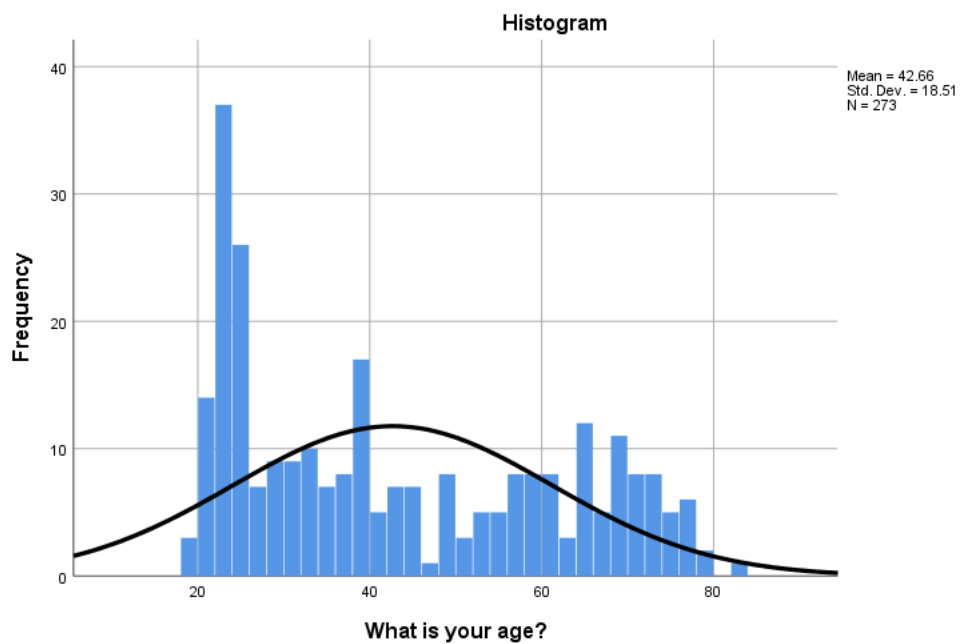


Figure 3.1: Age distribution of participants

Regarding familiarity with chatbots, nearly half of the respondents (48.7%) are quite familiar with them, while 28.2% have limited experience. A smaller segment is extremely familiar (14.3%), and 8.8% are not familiar at all (see figure 3.2).

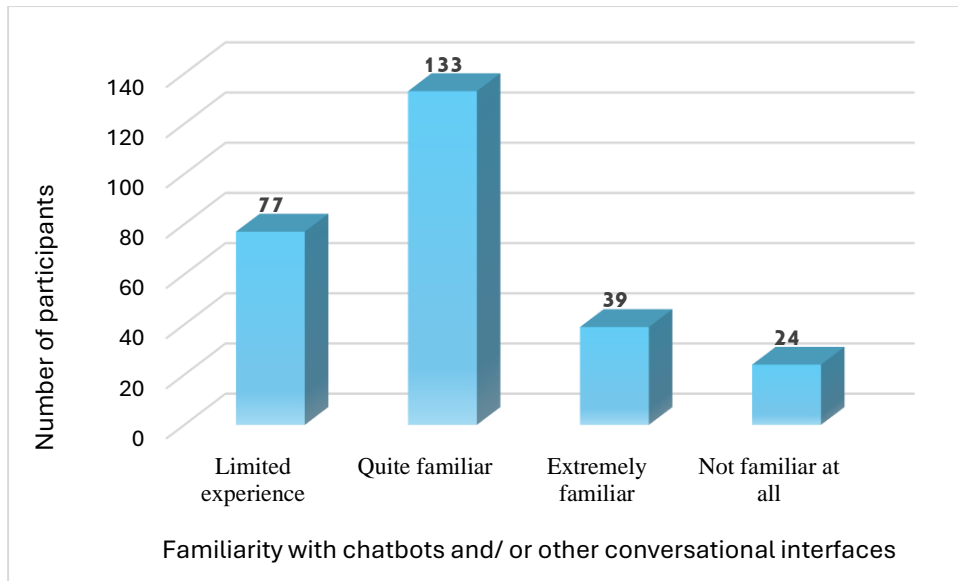


Figure 3.2: Familiarity with chatbots and/or other conversational interfaces

Usage of chatbots is relatively common, with 41.4% of respondents having definitely used them before and 39.6% using them sometimes. A minority has almost never used chatbots (9.2%), and 8.4% have definitely not used them (see figure 3.3).

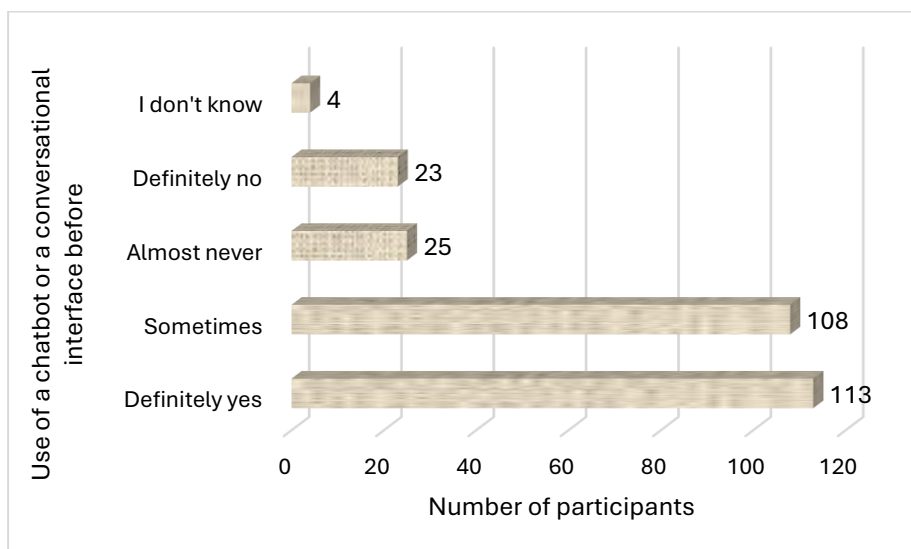


Figure 3.3: Use of a chatbot or a conversational interface before

Materials

The primary data collection instrument in this study was a scenario-based survey, presenting participants with four fictional conversations in a 2x2 full factorial design. The fictional conversations were with a banking AI chatbot designed to exhibit different levels of transparency by varying explainability (low/high) and accountability (low/high). To simulate closest AI chatbots, which are mainly powered by OpenAI's ChatGPT nowadays, the conversations were generated and refined through ChatGPT prompts. The detailed transcript of the conversations and screenshots of the interface are visible in Appendix B, and the prompts to generate the conversations are displayed in Appendix C.

Scenario 1 was for low accountability and explainability. In this scenario:

- Explainability is low because the chatbot fails to provide clear, concise explanations for its recommendations, actions, or the procedures to follow when an error is detected. The user is left without understanding why the error occurred or how the chatbot's transaction processing works.

- Accountability is low as the chatbot does not offer mechanisms within the chat to address or resolve the user's issue directly. It deflects responsibility to other channels and does not facilitate immediate feedback or corrective actions, nor does it empower the user to rectify issues through the chat interface.

Scenario 2 aimed at creating a high accountability and low explainability situation. In this version:

- High accountability is demonstrated by the chatbot taking immediate action to resolve the transaction error, providing options for direct feedback, and assuring the user that preventive measures are being reviewed.

- However, low explainability persists as the chatbot still fails to offer clear, detailed explanations of the processes involved or why the error occurred initially, not fully clarifying the underlying system operations or decision-making criteria.

Scenario 3 was for low accountability and high explainability. In this scenario:

- High explainability is demonstrated through the chatbot's detailed explanations of how errors in transactions might occur, the importance of double-checking details, and the processes involved in managing transactions.

- However, low accountability is evident as the chatbot lacks the authority or means to take corrective actions directly, pushing the user to seek resolution through other channels, thus not providing a direct resolution or feedback mechanism within the chat itself.

Finally, *scenario 4* was designed to reflect high accountability and high explainability. In this scenario:

- High Explainability is demonstrated through the chatbot's detailed explanations of how errors in transactions might occur, the importance of double-checking details, and the processes involved in managing transactions.

- High accountability is demonstrated by the chatbot taking immediate action to resolve the transaction error, providing options for direct feedback, and assuring the user that preventive measures are being reviewed.

Perceived trust

Perceived trust will be evaluated using the Chat Usability Scale. According to Borsci et al (2022, p.106), from an end-user perspective it is easier to assess 'trust' in a CRM chatbot interaction by assessing the bot's capacity to provide information and helping to attain a goal (i.e. the credibility of information) instead of by assessing trustworthiness as a general and unspecified sense of trust. Factor A3 from BUS-15 (4 items) is designed to assess perceived credibility which is also defined as trustworthiness by its authors. Each item will be assessed through a five-point Likert scale from 1 ('Strongly Disagree') to 5 ('Strongly Agree'). The perceived trust scale, with four items (PT1, PT2, PT3, PT4), demonstrated good reliability with a Cronbach's Alpha of 0.872.

Perceived expertise

Perceived expertise was gauged using the Generalized Expertise Measure (GEM) developed by Germain and Tejada (2012, p.221). This scale assesses the degree to which users perceive the chatbot as knowledgeable and competent. Divided in objective expertise and subjective expertise, we will only use the objective expertise items for the purpose of this study as the subjective items are phrased and directed at human traits and characteristics not applying in the context of the fictional discussions. The retained scale concerns six items, measured on a 7-point Likert-scale. Reliability of this subscale has an alpha of .92.

Behavioral intention to use of financial services

Lastly, behavioral intention to use of financial services was measured using the behavioral intention scale developed by Kaczmarek et al. (2014, p.90), which has shown high consistency through its alpha coefficient of .96. This scale evaluates the likelihood that users will engage in certain behaviors, such as engaging in financial services, based on their chatbot interaction. The scale has 3 items, and all can be retained for our study (BI_1, BI_2, BI_3) to assess the. The concise yet reliable aspect of this scale is important in optimizing the duration of the survey while ensuring a rigorous data collection and analysis.

Overview of scales

Overall, these results confirm that the items within each scale consistently measure their respective constructs (table 3.2).

Table 3.2

Reliability of the scales

Variables	Cronbach's Alpha	N of Items
Perceived Trust	0.872	4
Perceived Expertise	0.919	6
Behavioral Intention	0.962	3

The Cronbach's alpha values indicate that each item contributes to the scale's internal consistency, with values ranging from 0.714 to 0.775 if any single item is deleted, maintaining a high level of overall reliability (see table 3.3).

Table 3.3

Item-Total Statistics

	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item- Total Correlation	Cronbach's Alpha if Item Deleted
Perceived trust	70.08	221.442	0.74	0.714
Perceived expertise	56.94	113.47	0.75	0.749
Behavioral intention	74.03	198.242	0.568	0.742

Data quality control

Before the full-scale deployment of the survey, a pilot test was conducted with a small group of participants to ensure the clarity of scenario descriptions and survey items, test the reliability of the scales, and identify any technical issues or ambiguities in the survey flow.

Factor Analysis:

Perceived Trust

A principal component analysis (PCA) was conducted to examine the underlying structure of the perceived trust scale. The analysis was performed using SPSS Version 29 on a sample of 273 participants. The dataset comprised four items measuring perceived expertise.

The adequacy of the sample for PCA was confirmed by the Kaiser-Meyer-Olkin measure of sampling adequacy, $KMO = .82$, and Bartlett's Test of Sphericity, $\chi^2(6) = 552.61$, $p < .001$, indicating that the correlations between items were adequate for PCA. Based on the eigenvalue criterion (eigenvalues greater than 1) and inspection of the scree plot, one

component was retained, explaining 72.5% of the total variance. Table 3.4 presents the loadings of each item on the first principal component. All items showed strong loadings, ranging from .77 to .87, indicating that they are all related to a single underlying construct.

Table 3.4

Perceived trust – Component loading

Items	Component loading
Perceived_trust_1	.861
Perceived_trust_2	.879
Perceived_trust_3	.778
Perceived_trust_4	.877

GEM / General Expertise Measurement

A principal component analysis (PCA) was conducted to examine the underlying structure of the General Expertise Measurement scale. The analysis was performed using SPSS Version 29 on a sample of 273 participants. The dataset comprised six items measuring perceived expertise.

The adequacy of the sample for PCA was confirmed by the Kaiser-Meyer-Olkin measure of sampling adequacy, $KMO = .88$, and Bartlett's Test of Sphericity, $\chi^2(15) = 1165.46$, $p < .001$, indicating that the correlations between items were sufficiently large for PCA. Based on the eigenvalue criterion (eigenvalues greater than 1) and inspection of the scree plot, one component was retained, explaining 71.5% of the total variance. Table 3.5 presents the loadings of each item on the first principal component. All items showed strong loadings, ranging from .76 to .88, indicating that they are all related to a single underlying construct.

Table 3.5

Perceived expertise – Component loading

Items	Component loading
GEM_1	.783
GEM_2	.882
GEM_3	.815
GEM_4	.789
GEM_5	.833
GEM_6	.763

Behavioral intention to use financial services

A principal component analysis (PCA) was conducted to examine the underlying structure of the Behavioral Intention scale. The analysis was performed using SPSS Version 29 on a sample of 273 participants. The dataset comprised three items measuring behavioral intention to use financial services.

The adequacy of the sample for PCA was confirmed by the Kaiser-Meyer-Olkin measure of sampling adequacy, $KMO = .77$, and Bartlett's Test of Sphericity, $\chi^2(3) = 967.14$, $p < .001$, indicating that the correlations between items were sufficiently large for PCA. Based on the eigenvalue criterion (eigenvalues greater than 1) and inspection of the scree plot, one component was retained, explaining 92.99% of the total variance. Table 3.6 presents the loadings of each item on the first principal component. All three items showed especially strong loadings, ranging from .93 to .97, indicating that they are all strongly related to a their single underlying construct of behavioral intention to use financial services.

Table 3.6

Behavioral intention to use financial services – Component loading

Items	Component loading
BI_Use_Fin_Ser_1	.953
BI_Use_Fin_Ser_2	.970
BI_Use_Fin_Ser_3	.970

Procedure

The procedure for this study was designed to systematically examine the effects of chatbot transparency on user perceptions and behavioral intentions. A complete transcript of the questionnaire presented to participants is visible under Appendix A.

Participants were first presented with an informed consent form, which outlined the purpose of the study, the voluntary nature of participation, and assurances regarding the confidentiality of their responses. Upon agreeing to participate, they proceeded to the survey.

They were then asked to provide basic demographic information, including their age, gender, education level, and familiarity with chatbots. This information was collected to ensure a diverse sample and to analyze how demographic factors might influence perceptions of chatbot interactions.

Following the demographic questions, participants were randomly assigned to one of four chatbot scenarios. Each scenario was carefully crafted to simulate a conversation between a user and a banking chatbot, varying in levels of explainability and accountability. The scenarios were designed to last between one and two minutes, depending on the complexity and length of the interaction. The four scenarios included combinations of low and high explainability and accountability to examine their individual and combined effects on user perceptions.

The instructions given to participants were consistent across all scenarios. Participants were asked to imagine themselves as the user in the conversation and to carefully read the entire dialogue. They were instructed: “Please scroll down the following discussion between a chatbot and a user to read it. Imagine you are the user and this was the discussion you had.

Once you have finished, you can continue to the next question.” This approach aimed to immerse participants in the scenario and to ensure that their responses were based on a realistic simulation of chatbot interaction.

After reading the assigned scenario, participants were prompted to answer a series of survey questions designed to measure their perceptions of the chatbot's transparency, trustworthiness, expertise, and their behavioral intentions. After this they also had to complete two manipulation checks for explainability and accountability. These questions were presented in a structured format, utilizing Likert scales to capture the degree of agreement or disagreement with various statements about the chatbot's performance.

Data collection was conducted entirely online through the Qualtrics platform, which facilitated the randomization of scenarios, secure data handling, and efficient management of participant responses. The platform's features allowed for seamless transition between survey sections and ensured that each participant experienced a controlled and unbiased survey flow.

The study concluded with a debriefing section, where participants were thanked for their participation and provided with additional information about the study's objectives. This included an assurance that their data would be anonymized and used solely for research purposes.

Ethical Considerations

In accordance with ethical research practices, participants were provided with an informed consent form at the beginning of the survey. This form explained the purpose of the study, the voluntary nature of participation, and the confidentiality of their responses. Participants were assured that their data would be anonymized and used solely for research purposes. They were also informed of their right to withdraw from the study at any time without any penalty. Informing respondents about the true purpose of the study is an essential component of ethical research practice because it ensures that participants can make an informed decision about their involvement. This transparency respects the autonomy and dignity of the participants, as they are fully aware of what their participation entails and can thus provide genuine consent (Resnik, 2020, p. 45). Additionally, ethical guidelines, such as those outlined by the American Psychological Association, emphasize the importance of

honesty and transparency in research to maintain the integrity of the research process and foster trust between researchers and participants (APA, 2020, p. 19).

Results:

The relationship between transparency of AI chatbots and a user's behavioral intention to use financial services.

A univariate analysis of variance (ANOVA) was conducted to examine the effect of explainability and accountability levels of a chatbot on users' intention to use financial services. Descriptive statistics for the use of financial services across the levels of explainability and accountability are presented in Table 1. The results of the two-way ANOVA indicated no significant main effect of explainability on the use of financial services, $F(1, 269) = 0.156, p = .693, \eta^2 = .001$. There was also no significant main effect of accountability on the use of financial services, $F(1, 269) = 1.341, p = .248, \eta^2 = .005$. Furthermore, the interaction effect between explainability and accountability was not significant, $F(1, 269) = 1.806, p = .180, \eta^2 = .007$. These results suggest that neither explainability, accountability, nor their interaction had a significant impact on the use of financial services."

This indicates that variations in explainability and accountability, as defined in this study, do not seem to directly influence how participants intend to engage with financial services. An overview of the analysis' results is provided in table 4.1.

Table 4.1

Univariate Analysis of Variance (DV: Behavioral intention / IV: Explainability/Accountability)

Explainability condition	Accountability condition	Mean	SD	SE	95% CI	
					Lower	Upper
Low	Low	139	4.005	0.114	3.721	4.289
Low	High	134	4.086	0.147	3.797	4.375
High	Low	138	3.926	0.145	3.641	4.211
High	High	135	4.164	0.146	3.876	4.453

The results suggest that while the manipulations of explainability and accountability were successful, they did not significantly influence participants' behavioral intention to use financial services. Therefore, the hypothesis that there is a significant positive relationship between the explainability and accountability of AI chatbots and a user's behavioral intention to use financial services (H1) was not supported.

Perceived trust as a moderator of the relationship between transparency of AI chatbots and intention to use financial services.

A moderation analysis was conducted to examine whether perceived trust moderates the relationship between transparency (operationalized by accountability and explainability) of AI chatbots and behavioral intention to use financial services. The overall regression model was found to be significant, $F(3, 269) = 20.057, p < .001$, with an $R^2 = .183$, indicating that approximately 18.3% of the variance in behavioral intention to use financial services was explained by the predictors.

Table 4.2 presents the regression coefficients, standard errors, t-values, and p-values for each predictor in the model. The main effect of perceived trust on behavioral intention was significant, $B = .896, SE = .117, t = 7.645, p < .001$. However, the main effects of transparency through explainability and accountability were in both cases not significant.

Table 4.2

Moderation analysis (DV: Behavioral intention / IV: Accountability, explainability / Moderator: Perceived trust)

Predictor	B	SE	β	t	p
Constant	4.922	.210		23.334	<.001
Explainability (EXP_COND)	-.090	.188	-.026	-.476	.635
Accountability (ACC_COND)	-.147	.194	-.043	-.758	.449
Perceived trust (PT_CENT)	.896	.117	.439	7.645	<.001

These results point out that perceived trust significantly predicts behavioral intention to use financial services. However, the results also indicate that the main effects of explainability and accountability were not found to be significant in either case. This ultimately means that a moderation effect of perceived trust on the relationship between transparency and behavioral intention could not be confirmed with the respondent's data, rejecting hypothesis H2.

Perceived expertise as a moderator of the relationship between transparency of AI chatbots and behavioral intention in the banking industry.

A moderation analysis was conducted to examine whether perceived trust moderates the relationship between transparency (operationalized by accountability and explainability) of AI chatbots and behavioral intention to use financial services. The overall regression model (Model 2) was significant, $F(3, 269) = 53.148, p < .001$, with an $R^2 = .372$. The results of the ANOVA indicated a significant effect of the predictors on the dependent variable, behavioral intention, $F(3, 269) = 53.148, p < .001$.

Table 4.3 presents the regression coefficients, standard errors, t-values, and p-values for each predictor in the model. The main effect of perceived expertise on behavioral intention was significant, $B = .801, SE = .064, t = 12.538, p < .001$. However, the main effects of explainability and accountability were not significant.

Table 4.3

Moderation analysis (DV: Behavioral intention / IV: Accountability, explainability / Moderator: Perceived expertise)

Predictor	B	SE	β	t	p
Constant	3.614	.143		25.346	<.001
Explainability (EXP_COND)	-.295	.167	-.087	-1.767	.078
Accountability (ACC_COND)	-.219	.168	-.065	-1.304	.193
Perceived expertise (PE_CENT)	.801	.064	.630	12.538	<.001

These results show that perceived expertise significantly predicts behavioral intention to use financial services. Nevertheless, the results also indicate that the main effects of explainability and accountability were not found to be significant in either case again. This means that similarly to the previous hypothesis, a moderation effect of perceived expertise on the relationship between transparency and behavioral intention could not be confirmed with the respondent's data. Thus, our H3 is rejected.

Relationship between behavioral intention to use financial services, perceived trust and perceived expertise.

A regression analysis was conducted to examine whether the intention to use financial services is highest when both perceived trust and perceived expertise are high. The predictors included perceived trust (PT_COMP), perceived expertise (PE_COMP), and their interaction term (PT_PE_INT). The overall regression model was significant, $F(3, 269) = 51.517$, $p < .001$, with an $R^2 = .365$, indicating that approximately 36.5% of the variance in behavioral intention to use financial services was explained by the predictors. The results of the ANOVA indicated a significant effect of the predictors on the dependent variable, behavioral intention, $F(3, 269) = 51.517$, $p < .001$.

Table 4.4 presents the regression coefficients, standard errors, t-values, and p-values for each predictor in the model. The main effect of perceived trust on behavioral intention was not significant, $b = -.171$, $SE = .316$, $t = -.540$, $p = .590$. The main effect of perceived expertise on behavioral intention was not significant, $b = .473$, $SE = .275$, $t = 1.718$, $p = .087$. The interaction effect between perceived trust and perceived expertise was not significant, $b = .063$, $SE = .066$, $t = .952$, $p = .342$.

Table 4.4

Regression analysis (DV: BI_COMP / IV: PT_COMP, PE_COMP, PT_PE_IT)

Predictor	B	SE	β	t	p
(Constant)	1.141	1.158		.985	.325
PT * PE Interaction	.063	.066	.305	.952	.342
Perceived Expertise (PE_COMP)	.473	.275	.371	1.718	.087

Perceived Trust (PT_COMP)	-0.171	0.316	-0.084	-0.540	0.590
---------------------------	--------	-------	--------	--------	-------

The results interestingly suggest that, while the overall model significantly predicts behavioral intention to use financial services, each of the individual contributions of perceived trust, perceived expertise, and their interaction are not found to be statistically significant in any case. This indicates that the combination of high perceived trust and high perceived expertise does not significantly influence the behavioral intention to use financial services, nor does it when examined individually. Thus, our H4: “Intention to use financial services is highest, when both perceived trust and perceived expertise are high” is rejected.

Perceived expertise mediates the relationship between perceived trust and behavioral intention to use financial services.

To investigate the hypothesis that perceived trust mediates the relationship between perceived expertise and behavioral intention, a series of regression analyses were conducted following the guidelines by Baron and Kenny (1986, p.1176). The steps involved assessing the direct and indirect effects, and verifying the mediation effect using the Sobel test. The model of this mediation is presented in figure 3.4.

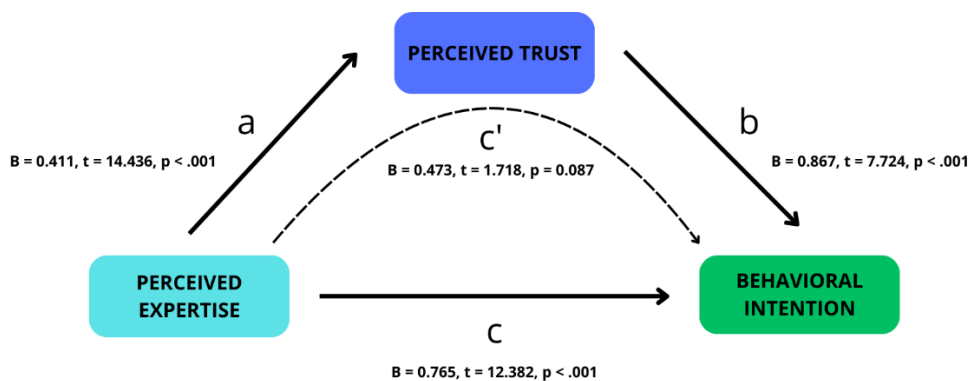


Figure 3.4 Mediation analysis model

Assumption checking

To ensure that the data met the assumptions required for mediation analysis, several diagnostic tests were conducted:

Linearity was assessed through scatterplots, which indicated that the relationships between perceived expertise and behavioral intention, as well as between perceived expertise and perceived trust, were linear. The normality of residuals was evaluated using histograms and Q-Q plots, which showed that the residuals were approximately normally distributed. However, the Shapiro-Wilk test indicated a slight deviation from normality ($p < .05$), which is often acceptable in large samples (Statistic = .982, $df = 273$, Sig. = .002). Homoscedasticity was examined by plotting the residuals against the predicted values. The scatterplot indicated that the residuals were randomly scattered around the horizontal axis without any discernible pattern, suggesting that the variance of the residuals was constant across all levels of the predicted values. The Durbin-Watson statistic was 2.098, indicating that the residuals were independent. Values close to 2 suggest that the assumption of independence of residuals is met, with no evidence of autocorrelation. Multicollinearity was assessed using the Variance Inflation Factor (VIF). The VIF values for perceived trust and perceived expertise were both 1.769, suggesting that multicollinearity was not a concern in this model.

Mediation analysis

1 - Direct effect of perceived expertise on behavioral intention (c path)

First, the direct effect of perceived expertise on behavioral intention was analyzed. The results indicated a significant positive relationship between perceived expertise and behavioral intention ($B = 0.765$, $SE = 0.062$, $t = 12.382$, $p < .001$). This suggests that users who perceive AI chatbots as highly expert are more likely to intend to use these chatbots for financial services. The high level of perceived expertise likely instills confidence in users, making them feel assured that the chatbot can competently handle their financial queries and tasks, thereby increasing their intention to use the chatbot. The results of this analysis are presented in table 4.5.

Table 4.5

Regression analysis (DV: BI_COMP / IV: PE_COMP)

Predictor	B	SE	β	t	p
(Constant)	0.317	0.312		1.017	0.310
Perceived Expertise	0.765	0.062	0.601	12.382	< .001

2 - Direct Effect of Perceived Expertise on Perceived Trust (a path)

Next, the effect of perceived expertise on perceived trust was analyzed. The findings revealed a significant positive relationship between perceived expertise and perceived trust ($B = 1.058$, $SE = 0.073$, $t = 14.436$, $p < .001$). This indicates that users who perceive the chatbot as having a high level of expertise are also more likely to trust the chatbot. Trust is a crucial factor in user interaction with AI systems, particularly in the financial services sector where accuracy and reliability are paramount. When users perceive a chatbot as knowledgeable and capable, their trust in the chatbot increases, which can facilitate a more positive and engaging user experience. The results of this analysis are presented in table 4.6.

Table 4.6

Regression analysis (DV: PT_COMP / IV: PE_COMP)

Predictor	B	SE	β	t	p
(Constant)	0.618	0.301		2.055	0.041
Perceived Expertise	1.058	0.073	0.659	14.436	< .001

3 - Direct Effect of Perceived Trust on Behavioral Intention (b path)

After this, the direct effect of perceived trust on behavioral intention was assessed. The results showed that perceived trust significantly predicts behavioral intention ($B = 0.867$, $SE = 0.112$, $t = 7.724$, $p < .001$). This finding underscores the importance of trust in influencing users' intention to use AI chatbots for financial services. Trust in the chatbot's abilities and integrity likely reduces users' apprehensions and concerns about using the

technology for sensitive financial transactions, thereby encouraging greater adoption and use. Results of this analysis are presented in table 4.7.

Table 4.7

Regression analysis (DV: BI_COMP / IV: PT_COMP)

Predictor	B	SE	β	t	p
(Constant)	0.556	0.461		1.207	0.228
Perceived Trust (PE_COMP)	0.867	0.112	0.425	7.724	< .001

4 - Direct effect of perceived expertise on behavioral intention while controlling for perceived trust (c' path)

Additionally, to evaluate the mediation effect, the direct effect of perceived expertise on behavioral intention was tested while controlling for perceived trust. The analysis revealed that when both perceived expertise and perceived trust were included in the model, perceived expertise remained a significant predictor of behavioral intention (B = 0.723, SE = 0.082, t = 8.789, p < .001), whereas perceived trust was not a significant predictor (B = 0.103, SE = 0.132, t = 0.779, p = .437). This indicates that the relationship between perceived expertise and behavioral intention is not mediated by perceived trust. The results from this analysis are presented in table 4.8.

Table 4.8

Regression analysis (DV: BI_COMP / IV: PT_COMP, PE_COMP)

Predictor	B	SE	β	t	p
(Constant)	0.110	0.410		0.267	0.789
Perceived Expertise	0.723	0.082	0.568	8.789	< .001
Perceived Trust	0.103	0.132	0.050	0.779	0.437

The results of this analysis offer a nuanced understanding of how perceived expertise and perceived trust independently influence users' intention to use AI chatbots in the context of financial services. Despite the significant roles both factors play, the mediation analysis reveals that perceived trust does not mediate the relationship between perceived expertise and behavioral intention.

Firstly, the significant positive relationship between perceived expertise and behavioral intention ($B = 0.765$, $SE = 0.062$, $t = 12.382$, $p < .001$) underscores the importance of users perceiving the chatbot as knowledgeable and competent. In the realm of financial services, where accuracy, reliability, and expert handling of information are paramount, users are more inclined to interact with and use chatbots that they perceive as possessing a high degree of expertise. This perception of expertise likely reassures users that the chatbot can manage complex financial queries and tasks, thereby increasing their confidence and willingness to use the service.

Moreover, the significant relationship between perceived expertise and perceived trust ($B = 1.058$, $SE = 0.073$, $t = 14.436$, $p < .001$) suggests that users are more likely to trust a chatbot they view as highly expert. Trust in AI systems, particularly in financial services, is a critical factor as users need to feel secure that the information provided by the chatbot is accurate and reliable. This trust reduces users' perceived risks associated with using the chatbot for sensitive financial transactions and thus can play a pivotal role in their overall user experience.

The direct effect of perceived trust on behavioral intention ($B = 0.867$, $SE = 0.112$, $t = 7.724$, $p < .001$) further emphasizes that trust is a significant predictor of users' intention to use AI chatbots. Trust in the chatbot ensures that users feel comfortable and secure while interacting with the chatbot, which can lead to higher levels of engagement and usage. This aligns with existing literature that highlights trust as a cornerstone in the adoption of AI technologies, especially in fields that deal with sensitive information like banking and finance.

However, the mediation analysis reveals that perceived trust does not mediate the relationship between perceived expertise and behavioral intention. When both perceived expertise and perceived trust are included in the model, perceived expertise remains a significant predictor of behavioral intention ($B = 0.723$, $SE = 0.082$, $t = 8.789$, $p < .001$),

while perceived trust does not ($B = 0.103$, $SE = 0.132$, $t = 0.779$, $p = .437$). This indicates that the impact of perceived expertise on users' intention to use the chatbot is direct and robust, not dependent on the level of trust users have in the chatbot.

This finding is particularly intriguing as it suggests that the perception of the chatbot's expertise alone is a strong enough driver of behavioral intention, independent of trust. In the context of AI chatbots in financial services, this could mean that users prioritize the chatbot's ability to handle complex and technical financial queries over their emotional or psychological comfort derived from trust. This highlights the critical role of perceived expertise in driving user adoption of AI chatbots. Users may be more focused on the functional capabilities and accuracy of the chatbot, which are directly tied to their perception of its expertise.

Overview of hypotheses:

An additional summary of the study's hypotheses and findings is given in Table 4.9 below.

Table 4.9

Summary of hypothesized relationships

Hypotheses	Results
H1: Higher transparency of AI chatbots leads to a higher intention to use financial services.	Not Supported
H2: Perceived trust moderates the relationship between transparency of AI chatbots and intention to use financial services.	Not Supported
H3: Perceived expertise moderates the relationship between transparency of AI chatbots and behavioral intention in the banking industry	Not Supported
H4: Behavioral intention to use financial services is highest, when both perceived trust and perceived expertise are high.	Not Supported
H5: Perceived trust mediates the relationship between perceived expertise and behavioral intention	Not Supported

Conclusion

Overview of research findings

This study aimed to explore various aspects of how AI chatbot transparency influences user behavior in the context of financial services, with a particular focus on the roles of perceived trust and perceived expertise. Several relationships were hypothesized and tested through a series of statistical analyses, including univariate ANOVA, regression analysis, and moderation analysis.

Hypothesis 1: Transparency of AI chatbots and behavioral intention

The first hypothesis posited that higher transparency (operationalized through explainability and accountability) of AI chatbots would lead to a higher intention to use financial services. The two-way ANOVA results indicated no significant main effects for either explainability ($F(1, 269) = 0.156, p = .693$) or accountability ($F(1, 269) = 1.341, p = .248$). Additionally, the interaction effect between explainability and accountability was also not significant ($F(1, 269) = 1.806, p = .180$). These findings suggest that the variations in explainability and accountability did not significantly influence participants' behavioral intention to use financial services, thus not supporting Hypothesis 1.

Hypothesis 2: Perceived trust as a moderator

The second hypothesis suggested that perceived trust moderates the relationship between AI chatbot transparency and behavioral intention. The overall regression model was significant ($F(3, 269) = 20.057, p < .001, R^2 = .183$), indicating that approximately 18.3% of the variance in behavioral intention was explained by the predictors. However, while perceived trust was a significant predictor ($B = .896, SE = .117, t = 7.645, p < .001$), the main effects of explainability and accountability were not significant, nor was the moderation effect confirmed. Consequently, Hypothesis 2 was not supported.

Hypothesis 3: Perceived expertise as a moderator

The third hypothesis proposed that perceived expertise moderates the relationship between AI chatbot transparency and behavioral intention. The regression model was significant ($F(3, 269) = 53.148, p < .001, R^2 = .372$), with perceived expertise significantly predicting behavioral intention ($B = .801, SE = .064, t = 12.538, p < .001$). However, the main effects of explainability and accountability remained non-significant, and the moderation effect was not supported, leading to the rejection of Hypothesis 3.

Hypothesis 4: Combined effect of perceived trust and expertise

Hypothesis 4 explored whether behavioral intention to use financial services is highest when both perceived trust and perceived expertise are high. The regression analysis, while indicating a significant overall model ($F(3, 269) = 51.517, p < .001, R^2 = .365$), showed that the individual contributions of perceived trust, perceived expertise, and their interaction were not statistically significant. Thus, the hypothesis that the highest behavioral intention occurs when both perceived trust and perceived expertise are high was not supported.

Hypothesis 5: Mediation effect of perceived trust

Finally, hypothesis 5 investigated a possible mediation effect of perceived trust between perceived expertise and intention to use financial services. The results suggest that while both perceived expertise and perceived trust are important factors independently influencing users' intention to use AI chatbots for financial services, perceived trust does not mediate the relationship between perceived expertise and behavioral intention. In other words, users' intention to use the banking chatbot is directly influenced by their perception of the chatbot's expertise, irrespective of their level of trust in the chatbot. This highlights the critical role of perceived expertise in driving user adoption of AI chatbots in the financial sector.

Discussion

The findings from this study provide significant insights into the dynamics of user interaction with AI chatbots in the financial services sector. This section discusses the implications of these findings, relating them back to the original theoretical framework and existing literature.

Transparency, Perceived Trust, and Behavioral Intention

This research aimed to examine the impact of transparency—operationalized through explainability and accountability—on users' behavioral intention to use financial services facilitated by AI chatbots. Contrary to expectations, the results indicated that variations in explainability and accountability did not significantly influence participants' behavioral intentions. This finding challenges the widely held belief that increased transparency directly correlates with higher user engagement and adoption rates (Walmsley, 2021, p.586).

Despite the lack of a direct effect, perceived trust emerged as a significant predictor of behavioral intention, which aligns with existing research emphasizing trust as a critical factor in the acceptance of AI technologies (Yang & Wibowo, 2022, p.2060). The significant relationship between perceived expertise and perceived trust ($B = 1.058$, $SE = 0.073$, $t = 14.436$, $p < .001$) underscores the importance of users perceiving the chatbot as knowledgeable and competent. This perception not only fosters trust but also enhances users' confidence in the chatbot's ability to manage complex financial queries, thus encouraging engagement (Choung et al., 2022, p.9).

Perceived Expertise, Perceived Trust, and Behavioral Intention

The mediation analysis revealed that perceived trust does not mediate the relationship between perceived expertise and behavioral intention. While perceived expertise was a strong predictor of behavioral intention ($B = 0.765$, $SE = 0.062$, $t = 12.382$, $p < .001$), its influence remained robust even when controlling for perceived trust. This finding suggests that users' behavioral intention to use financial services through AI chatbots is driven more by their perception of the chatbot's expertise rather than trust alone. This could indicate that in the context of financial services, users prioritize the chatbot's functional capabilities and the

accuracy of its responses over their emotional comfort derived from trust (Mori & Du, 2023, p.3).

The direct effect of perceived trust on behavioral intention ($B = 0.867$, $SE = 0.112$, $t = 7.724$, $p < .001$) reinforces the critical role of trust in user engagement with AI systems. Trust reduces perceived risks associated with using AI for sensitive tasks, such as financial transactions, thereby facilitating user adoption (Lukyanenko et al., 2022, p.2000). However, the non-significance of perceived trust as a mediator highlights a potential divergence between trust and expertise in driving user behavior. Users may inherently trust AI systems that exhibit high expertise without requiring additional trust-building measures (Yang & Wibowo, 2022, p.2062).

The Combined Effect of Perceived Trust and Expertise

Hypothesis 4, which proposed that behavioral intention would be highest when both perceived trust and perceived expertise are high, was not supported by the data. The interaction term between perceived trust and perceived expertise was not significant, suggesting that these factors operate independently rather than synergistically to influence behavioral intention. This finding challenges the assumption that enhancing both trust and expertise simultaneously would have a compounded effect on user engagement (Wanner et al., 2022, pp.2095-2096).

Academical and professional implications

Academically

The findings of this study can hold several academic implications, particularly within the context of AI chatbot research in the banking sector. By examining the roles of transparency, perceived trust, perceived expertise, and behavioral intention, this research contributes to a deeper understanding of the psychological and behavioral mechanisms at play in user interactions with AI systems.

Transparency and user behavior

The study's findings underscore the complexity of transparency's impact on user behavior. Although higher levels of explainability and accountability did not directly influence behavioral intention, they play a critical role in shaping perceived trust and expertise. This aligns with the theoretical framework posited by Walmsley (2021), which emphasizes that transparency is a multifaceted concept involving clear communication of an AI system's operations, decision-making processes, and underlying algorithms (2021, p. 586). Thus, transparency should be seen as a foundational element that indirectly shapes user behavior through its impact on mediating factors such as trust and perceived expertise. The extent of the scope of variables at play here is what needs to be further investigated. This would provide a more comprehensive approach to understanding a user's behavioral intention in chatbot interactions.

Perceived trust and expertise as predictors of intention to use financial services

The study highlights that both perceived trust and perceived expertise independently predict users' intention to use AI chatbots for financial services. Perceived expertise emerged as a strong predictor of behavioral intention, indicating that users are more likely to engage with chatbots they view as knowledgeable and competent. Similarly, perceived trust significantly influences behavioral intention, emphasizing the importance of users' confidence in the chatbot's reliability and integrity. However, the lack of a mediation effect suggests that perceived trust and perceived expertise operate independently in shaping user intentions, with expertise having a direct impact on behavioral intention. This partly supports existing theories that trust is a crucial determinant in the acceptance of AI technologies (Yang & Wibowo, 2022, p. 2060). This finding also aligns with prior research indicating that trust enhances user engagement and willingness to adopt AI systems (Wanner et al., 2022, p. 2095). However, the study extends this understanding by demonstrating that perceived expertise can enhance behavioral intention, as well as trust, which in turn increases behavioral intention. This suggests a cascading effect where expertise builds trust, and trust drives behavior, but without a mediating effect of trust between expertise and behavioral intention. These results and findings provide a more integrated view of how these constructs interact in the context of a banking chatbot.

Implications for theoretical models

Finally, the findings contribute to a broader theoretical framework of AI, chatbots, and more largely technology acceptance models by integrating elements of AI transparency, trust, and expertise. Traditional models can be enriched by incorporating these factors, particularly in the context of AI-driven interactions. By doing so, researchers can develop more comprehensive models that better predict user acceptance and usage behaviors in AI contexts (Gatzioufa & Saprikis, 2022, p.5).

Professionally

The results of this study also hold interesting implications for the professional world. Since AI and chatbots now span across different industries, the scope of workers that could benefit from these insights is very large. They can be related, but not limited, to the development, design, operationalization, marketing or management of AI-user interactions and chatbots. By understanding the factors that influence user acceptance and use of AI chatbots, banking institutions can tailor their chatbot design and implementation strategies to enhance user trust, perceived expertise, and ultimately, behavioral intention to use the chatbot for financial services. The study's findings offer several concrete recommendations for the banking sector which will be discussed in the following section.

Trust as a central factor

Firstly, enhancing perceived trust is a central point of importance, as it appears to be a driver of behavioral intention. Chatbots should provide clear, accurate, and honest information to build trust. For this, they need to achieve transparency in their responses, such as disclosing when a user is interacting with an AI and providing access to human support when needed. Also, emphasizing strong data privacy and security measures within the chatbot interactions can reassure users about the safety of their personal information, ultimately increasing trust. Ensuring that chatbots provide consistent and reliable responses can also help in building and maintaining trust over time.

Perceive expertise driving trust

Improving perceived expertise is another important factor. Developing chatbots with advanced AI capabilities, that can handle complex queries and provide expert-level information, can help enhance the perceived expertise of the chatbot. Additionally, continuously updating the chatbot's knowledge base and training it to handle a wide range of financial queries can make users perceive it as more knowledgeable and capable. Lastly, implementing context-aware responses, where the bot can remember past interactions and use them to provide more personalized and relevant information, could improve perceived expertise, which would in return affect perceived trust and intention to use financial services.

Implementing these recommendations can help banking institutions leverage AI chatbots more effectively, enhancing user trust and perceived expertise, which are crucial for driving user engagement and behavioral intention to use financial services. By focusing on these key areas, banks can create more meaningful and impactful interactions with their customers, ultimately leading to increased satisfaction and loyalty.

Limitations and future research

Finally, while the study's theoretical framework provided a robust foundation for exploring its central concepts, it is important to acknowledge a limitation in its approach of AI transparency. As discussed by Balasubramaniam et al. (2023, p.2) the complex concepts studied here are often interrelated with other transparency or acceptance-related concepts, such as privacy or reliability, and behave differently in the presence or absence of other influencing factors.

Given the study's findings, future research should further explore the pathways through which transparency affects user behavior. For instance, this could be achieved by manipulating the variables to make the difference in chatbot responses stronger, and/or making the interaction longer and covering more aspects. This could provide a more comprehensive environment for participants and maybe with the addition of new concepts to

explainability and accountability, the perception of transparency, the behavioral intention, and the relationship between the two variables will produce different results.

Alternatively, longitudinal studies could also provide deeper insights into how trust and expertise develop over time with repeated interactions. Lastly, experimental studies that manipulate different aspects of transparency, such as varying levels of detail in explanations or different accountability mechanisms, could further help to understand the specific impact these concepts have on user perceptions and behaviors.

In summary, the study opens avenues for exploring other potential mediators and moderators in the relationship between AI transparency and user behavior. Factors such as privacy, perceived risk, user autonomy, and emotional responses to AI interactions could be investigated to provide a more holistic understanding of user engagement with AI systems.

Conclusion

Ultimately, this study was looking to answer the following research question: *“How does transparency of AI chatbots affect a user’s behavioral intention to use financial services through perceived trust and perceived expertise?”*. In definitive, it can be said that its findings contribute to the growing body of literature on AI chatbots in financial services by highlighting the distinct and critical roles of perceived expertise and trust. While transparency in the form of explainability and accountability did not directly influence behavioral intention, perceived expertise and trust emerged as pivotal factors, although the interplay and relationship of the two still remains to be investigated in further research.

Literature

- Abdul, A., Vermeulen, J., Wang, D., Lim, B. Y., & Kankanhalli, M. (2018). Trends and trajectories for explainable, accountable and intelligible systems: An hci research agenda. *Proceedings of the 2018 CHI conference on human factors in computing systems*, 1-18. <https://doi.org/10.1145/3173574.3174156>
- Abu-Rasheed, H., Abdulsalam, M. H., Weber, C., & Fathi, M. (2024). Supporting Student Decisions on Learning Recommendations: An LLM-Based Chatbot with Knowledge Graph Contextualization for Conversational Explainability and Mentoring. <https://doi.org/10.48550/arXiv.2401.08517>
- American Psychological Association. (2020). *Publication manual of the American Psychological Association: The official guide to APA style* (7th ed.).
- Balasubramaniam, N., Kauppinen, M., Rannisto, A., Hiekkänen, K., & Kujala, S. (2023). Transparency and explainability of AI systems: From ethical guidelines to requirements. *Information and Software Technology*, 159, <https://doi.org/10.1016/j.infsof.2023.107197>
- Baron, R. M., & Kenny, D. A. (1986). The moderator–mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology*, 51(6), 1173–1182. <https://doi.org/10.1037/0022-3514.51.6.1173>
- Borsci, S., Malizia, A., Schmettow, M., Van Der Velde, F., Tariverdiyeva, G., Balaji, D., & Chamberlain, A. (2022). The Chatbot Usability Scale: the design and pilot of a usability scale for interaction with AI-based conversational agents. *Personal and Ubiquitous Computing*, 26, 95-119. <https://doi.org/10.1007/s00779-021-01582-9>
- Buijsman, S. (2024). Transparency for AI systems: a value-based approach. *Ethics and Information Technology*, 26(2), 34. <https://doi.org/10.1007/s10676-024-09770-w>
- Burrell, J. (2016). How the machine ‘thinks’: Understanding opacity in machine learning algorithms. *Big data & society*, 3(1), <https://doi.org/10.1177/205395171562251>
- Chazette, L., Brunotte, W., & Speith, T. (2021). Exploring explainability: a definition, a model, and a knowledge catalogue. *2021 IEEE 29th international requirements*

engineering conference, 197-208. IEEE.

<https://doi.org/10.1109/RE51729.2021.00025>

Choung, H., David, P., & Ross, A. (2022). Trust in AI and its role in the acceptance of AI technologies. *International Journal of Human–Computer Interaction*, 1–13.

<https://doi.org/10.1080/10447318.2022.2050543>

Creswell, J. W., & Creswell, J. D. (2017). *Research design: Qualitative, quantitative, and mixed methods approaches*. Sage publications.

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>

Diakopoulos, N. (2015). Algorithmic accountability: Journalistic investigation of computational power structures. *Digital journalism*, 3(3), 398-415.

<https://doi.org/10.1080/21670811.2014.976411>

Diakopoulos, N. (2016). Accountability in algorithmic decision making. *Communications of the ACM*, 59(2), 56-62. <https://doi.org/10.1145/2844110>

Diakopoulos, N., & Koliska, M. (2017). Algorithmic transparency in the news media. *Digital journalism*, 5(7), 809-828. <https://doi.org/10.1080/21670811.2016.1208053>

Field, A. (2013). *Discovering statistics using IBM SPSS statistics*. Sage publications.

Gatzioufa, P., & Saprikis, V. (2022). A literature review on users' behavioral intention toward chatbots' adoption. *Applied Computing and Informatics*, (ahead-of-print).

<https://doi.org/10.1108/ACI-01-2022-0021>

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>

Germain, M. L., & Tejada, M. J. (2012). A preliminary exploration on the measurement of expertise: An initial development of a psychometric scale. *Human Resource Development Quarterly*, 23(2), 203-232. <https://doi.org/10.1002/hrdq.21134>

- Grennan, L., Kremer, A., Singla, A., & Zipparo, P. (2022, September 29). Why businesses need explainable AI. *McKinsey*.
<https://www.mckinsey.com/capabilities/quantumblack/our-insights/why-businesses-need-explainable-ai-and-how-to-deliver-it>
- Jafarkarimi, H., Sim, A. T. H., Saadatdoost, R., & Hee, J. M. (2016). Designing a scenario-based questionnaire to assess behavioral intention in social networking sites' ethical dilemmas. In *Blurring the boundaries through digital innovation: Individual, organizational, and societal challenges* (pp. 145-159). Springer International Publishing. https://doi.org/10.1007/978-3-319-38974-5_12
- Kaczmarek, L. D., Goodman, F. R., Drażkowski, D., Kashdan, T. B., Połatyńska, K., & Komorek, J. (2014). Instructional support decreases desirability and initiation of a gratitude intervention. *Personality and Individual Differences*, 64, 89-93.
<https://dx.doi.org/10.1016/j.paid.2014.02.0220191-8869/>
- Kuhail, M. A., Alturki, N., Alramlawi, S., & Alhejori, K. (2023). Interacting with educational chatbots: A systematic review. *Education and Information Technologies*, 28(1), 973-1018. <https://doi.org/10.1007/s10639-022-11177-3>
- Lepri, B., Oliver, N., Letouzé, E., Pentland, A., & Vinck, P. (2018). Fair, transparent, and accountable algorithmic decision-making processes: The premise, the proposed solutions, and the open challenges. *Philosophy & Technology*, 31(4), 611-627.
<https://doi.org/10.1007/s13347-017-0279-x>
- Liao, Q. V., & Wortman Vaughan, J. (2024). AI Transparency in the Age of LLMs: A Human-Centered Research Roadmap. *Harvard Data Science Review*, (Special Issue 5). <https://doi.org/10.1162/99608f92.8036d03b>
- Liu, B. (2021). In AI we trust? Effects of agency locus and transparency on uncertainty reduction in human–AI interaction. *Journal of computer-mediated communication*, 26(6), 384-402. <https://doi.org/10.1093/jcmc/zmab013>
- 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., 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>
- 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>
- Mori, M., & Du, L. (2023). AI chatbots for banks: Evolving trends and critical issues. *The International Conference on Artificial Intelligence and Logistics Engineering*, 3-13. https://doi.org/10.1007/978-3-031-36115-9_1
- Novelli, C., Taddeo, M., & Floridi, L. (2023). Accountability in artificial intelligence: what it is and how it works. *AI & Society*, 1-12. <https://doi.org/10.1007/s00146-023-01635-y>
- Pavlou, P. A., & Fygenson, M. (2006). Understanding and predicting electronic commerce adoption: An extension of the theory of planned behavior. *MIS Quarterly*, 30(1), 115-143. <https://doi.org/10.2307/25148720>
- Resnik, D. B. (2018). *The ethics of research with human subjects: Protecting people, advancing science, promoting trust* (Vol. 74). Springer.
- Shin, D. (2020). User perceptions of algorithmic decisions in the personalized AI system: Perceptual evaluation of fairness, accountability, transparency, and explainability. *Journal of Broadcasting & Electronic Media*, 64(4), 541–565. <https://doi.org/10.1080/08838151.2020.1843357>
- Shin, D. (2021). Why does explainability matter in news analytic systems? Proposing explainable analytic journalism. *Journalism Studies*, 22(8), 1047–1065. <https://doi.org/10.1080/1461670X.2021.1916984>
- Shin, D., & Park, Y. J. (2019). Role of fairness, accountability, and transparency in algorithmic affordance. *Computers in Human Behavior*, 98, 277–284. <https://doi.org/10.1016/j.chb.2019.04.019>
- Walmsley, J. (2021). Artificial intelligence and the value of transparency. *AI & Society*, 36(2), 585-595. <https://doi.org/10.1007/s00146-020-01066-z>

- 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>
- Wu, W., Zhang, B., Li, S., & Liu, H. (2022). Exploring factors of the willingness to accept AI-assisted learning environments: an empirical investigation based on the UTAUT model and perceived risk theory. *Frontiers in psychology*, 13, 870777. <https://doi.org/10.3389/fpsyg.2022.870777>
- Yang, R., & Wibowo, S. (2022). User trust in artificial intelligence: A comprehensive conceptual framework. *Electronic Markets*, 32(4), 2053-2077. <https://doi.org/10.1007/s12525-022-00592-6>

Appendix A: Questionnaire

MASTER THESIS QUESTIONNAIRE

Survey Flow

EmbeddedData

assignmentIDValue will be set from Panel or URL.

participantIDValue will be set from Panel or URL.

Block: Default Question Block (1 Question)

BlockRandomizer: 1 - Evenly Present Elements

Standard: Fictional Chat 1 - Lo Exp / Lo Acc (2 Questions)

Standard: Fictional Chat 2 - Hi Exp / Lo Acc (2 Questions)

Standard: Fictional Chat 3 - Lo Exp / Hi Ac (2 Questions)

Standard: Fictional Chat 4 - Hi Exp / Hi Acc (2 Questions)

Standard: 1 - BUS15 (4 Questions)

Standard: GEM (1 Question)

Standard: BI Scale (1 Question)

Standard: Manipulation check (2 Questions)

Standard: Demographics (6 Questions)

Standard: Ending (1 Question)

Page Break

Start of Block: Default Question Block

Consent

Dear participant,

Thank you very much for participating in this research. This research is conducted for a master's thesis project for Media & Business at Erasmus University Rotterdam.

You will first be presented with a fictional discussion between a banking chatbot and a user. After this, you will have to answer questions on your impressions and perception of the discussion and of the chatbot.

Please be aware that your participation is completely voluntarily, meaning that you can quit at any time during your participation. Furthermore, your personal information will be kept strictly confidential, and the findings of this survey will be used solely for class purposes. Hence, your anonymity is guaranteed. Completing the survey takes approximately 5 minutes. If you have any questions during or after your participation, please feel free to contact me: Tasio Lopez--Heurtin (535928tl@student.eur.nl).

I understand the above and agree on participating in this research. (4)

End of Block: Default Question Block

Start of Block: Fictional Chat 1 - Lo Exp / Lo Acc



1 - Lo Exp / Lo Acc Please **scroll down** the following discussion between a chatbot and a user to read it. Imagine you are the user and this was the discussion you had. Once you have finished, you can continue to the next question.

Validation Scenario1 Click "Ok" when you are done reading and you can continue to the questionnaire.

Ok (1)

End of Block: Fictional Chat 1 - Lo Exp / Lo Acc

Start of Block: Fictional Chat 2 - Hi Exp / Lo Acc



2 - Hi Exp / Lo Acc Please **scroll down** the following discussion between a chatbot and a user to read it. Imagine you are the user and this was the discussion you had. Once you have finished, you can continue to the next question.

Validation Scenario2 Click "Ok" when you are done reading and you can continue to the questionnaire.

Ok (1)

End of Block: Fictional Chat 2 - Hi Exp / Lo Acc

Start of Block: Fictional Chat 3 - Lo Exp / Hi Ac



3 - Lo Exp / Hi Ac Please **scroll down** the following discussion between a chatbot and a user to read it. Imagine you are the user and this was the discussion you had. Once you have finished, you can continue to the next question.

Validation Scenario3 Click "Ok" when you are done reading and you can continue to the questionnaire.

Ok (1)

End of Block: Fictional Chat 3 - Lo Exp / Hi Ac

Start of Block: Fictional Chat 4 - Hi Exp / Hi Acc



4 - Hi Exp / Hi Acc Please **scroll down** the following discussion between a chatbot and a user to read it. Imagine you are the user and this was the discussion you had. Once you have finished, you can continue to the next question.

Validation Scenario4 Click "Ok" when you are done reading and you can continue to the questionnaire.

Ok (1)

End of Block: Fictional Chat 4 - Hi Exp / Hi Acc

1.1 Accessibility We would like to know your thoughts on this interaction between a banking chatbot and a user. Please complete the following statements according to your impressions, ranging from "Completely disagree" to "Completely agree".

	Completely disagree (1)	Somewhat disagree (2)	Neither agree nor disagree (3)	Somewhat agree (4)	Completely agree (5)
The chatbot function was easily detectable (2)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
It was easy to find the chatbot (3)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Page Break

1.2 Quality of func We would like to know your thoughts on this interaction between a banking chatbot and a user. Please complete the following statements according to your impressions, ranging from "Completely disagree" to "Completely agree".

	Completely disagree (1)	Somewhat disagree (2)	Neither agree nor disagree (3)	Somewhat agree (4)	Completely agree (5)
Communicating with the chatbot was clear. (1)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I was immediately made aware of what information the chatbot can give me. (3)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
The interaction with the chatbot felt like an ongoing conversation. (4)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
The chatbot was able to keep track of context. (5)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
The chatbot was able to make references to the website or service when appropriate. (6)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
The chatbot could handle situations in which the line of conversation was not clear. (7)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
The chatbot's responses were easy to understand. (8)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

1.3 Quality of info We would like to know your thoughts on this interaction between a banking chatbot and a user. Please complete the following statements according to your impressions, ranging from "Completely disagree" to "Completely agree".

	Completely disagree (1)	Somewhat disagree (2)	Neither agree nor disagree (3)	Somewhat agree (4)	Completely agree (5)
I find that the chatbot understands what I want and helps me achieve my goal. (1)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
The chatbot gives me the appropriate amount of information. (2)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
The chatbot only gives me the information I need. (3)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I feel like the chatbot's responses were accurate. (4)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Page Break

1.4 Privacy/Security We would like to know your thoughts on this interaction between a banking chatbot and a user. Please complete the following statements according to your impressions, ranging from "Completely disagree" to "Completely agree".

	Completely disagree (1)	Somewhat disagree (2)	Neither agree nor disagree (3)	Somewhat agree (4)	Completely agree (5)
I believe the chatbot informs me of any possible privacy issues. (1)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

 Page Break

End of Block: 1 - BUS15

Please continue and complete the following statements based on your impressions regarding the banking chatbot discussion.

Q2 - 55M

Strongly disagree (1) Disagree (2) Somewhat disagree (3) Neither agree nor disagree (4) Somewhat agree (5) Agree (6) Strongly agree (7)

This chatbot has knowledge that is specific to each request. (2)

This chatbot shows that they have the reasoning and information necessary to be an expert in its field. (banking advice) (8)

This chatbot has knowledge about its field (banking advice). (15)

This chatbot conducts research related to its field. (banking advice) (9)

This chatbot has the qualifications required to be an expert in its field. (banking advice) (10)



This chatbot has been trained in its area of expertise. (banking advice) (11)



End of Block: GEM

Start of Block: BI Scale



Q3 - BI scale Please continue and complete the following statements based on your impressions regarding the banking chatbot discussion.

	Strongly disagree (1)	Disagree (2)	Somewhat disagree (3)	Neither agree nor disagree (4)	Somewhat agree (5)	Agree (6)	Strongly agree (7)
I intend to use banking chatbots in the future. (2)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I intend to increase my use of banking chatbots in the future. (15)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I intend to use banking chatbots frequently in the future. (16)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

End of Block: BI Scale

Start of Block: Manipulation check

MC - Explainability Please indicate your impressions regarding the following statements.

	Strongly disagree (1)	Disagree (2)	Neither agree nor disagree (3)	Agree (4)	Strongly agree (5)
In the discussion, the chatbot is able to explain its actions and decisions in a way that is understandable to the user. This involves providing clear, concise, and understandable explanations for its recommendations or actions. (1)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Page Break

MC - Accountability Please indicate your impressions regarding the following statements.

	Strongly disagree (1)	Disagree (2)	Neither agree nor disagree (3)	Agree (4)	Strongly agree (5)
In the discussion, the chatbot is accountable for its actions and decisions. This means that there are mechanisms in place for users to provide feedback, report issues, and seek redress if the chatbot makes a mistake. (1)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

End of Block: Manipulation check


Start of Block: Demographics

Q4 - Age

Before finishing the survey, we would like to ask you a few questions about yourself.

What is your age?

15 20 26 31 36 42 47 53 58 63 69 74 79 85 90

Age ()	
--------	--

Q5 - Gender What gender do you identify with?

- Male (1)
- Female (2)
- Non-binary / third gender (3)
- Other (4)
- Prefer not to say (5)

Q6 - Education What is the highest educational level that you have **followed**? This can either be an education that you completed or one that you are or were previously enrolled in.

- Primary school (1)
 - Secondary school / high school (2)
 - Vocational degree after high school (3)
 - Bachelor's degree (4)
 - Master's degree (5)
 - PhD, MBA, or other equivalent (6)
 - Other, namely (7) _____
 - Prefer not to say (8)
-

Q7 - BC experience How familiar are you with chatbots and/or other conversational interfaces?

	I don't know (1)	Not familiar at all (8)	Limited experience (2)	Quite familiar (4)	Extremely familiar (5)
Choose an option (1)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

BC Experience 2 Have you used a chatbot or a conversational interface before?

	I don't know (3)	Definitely no (1)	Almost never (2)	Sometimes (4)	Definitely yes (5)
Choose an option (1)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

BC Experience 3 How often do you use chatbots weekly?

	I don't know (3)	Never (1)	Around once a week (2)	Every other day (4)	Daily (5)
Choose an option (1)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

End of Block: Demographics

Start of Block: Ending

You have now reached the end of the questionnaire. Thank you for your time and effort. Your help is highly appreciated! If you have questions or comments about this questionnaire, please list them below.

PLEASE PRESS THE BLUE ARROW BUTTON TO STORE ALL YOUR ANSWERS.

If you are taking this survey to earn rewards, you will have the links/codes on the next page.

End of Block: Ending

Appendix B: Stimuli / Fictional conversations

1. Low Explainability and low accountability

User: Hi there! I need to check my recent transactions and possibly transfer some money. Can you assist me with that?

Chatbot: Hello! I can show you your transactions.

User: Great, here's my account number: [account number]. Can you list my recent transactions?

Chatbot: Transactions from the last month are now displayed in your account section.

(The chatbot does not provide any detail on how to navigate to the account section or what specifically to look for.)

User: I noticed a charge for \$200 that should have been \$20. Can you fix this?

Chatbot: For issues with transactions, you need to contact our support team directly.

User: Why did this error happen?

Chatbot: Errors can occur. Please contact support to resolve transaction issues.

(The chatbot provides no explanation about the error or how the banking system processes transactions, showing low explainability.)

User: How can I prevent this in the future?

Chatbot: Always double-check the details before confirming a transaction.

User: What if I have other issues? Who can help me?

Chatbot: Our support team is available by phone or at our branches.

User: Is there a way to give feedback on this chat or report the issue for review?

Chatbot: You can use the feedback form on our website to report issues.

(The chatbot does not facilitate direct feedback or issue reporting through the chat interface, showing low accountability.)

User: Okay, I'll look into that. Thanks.

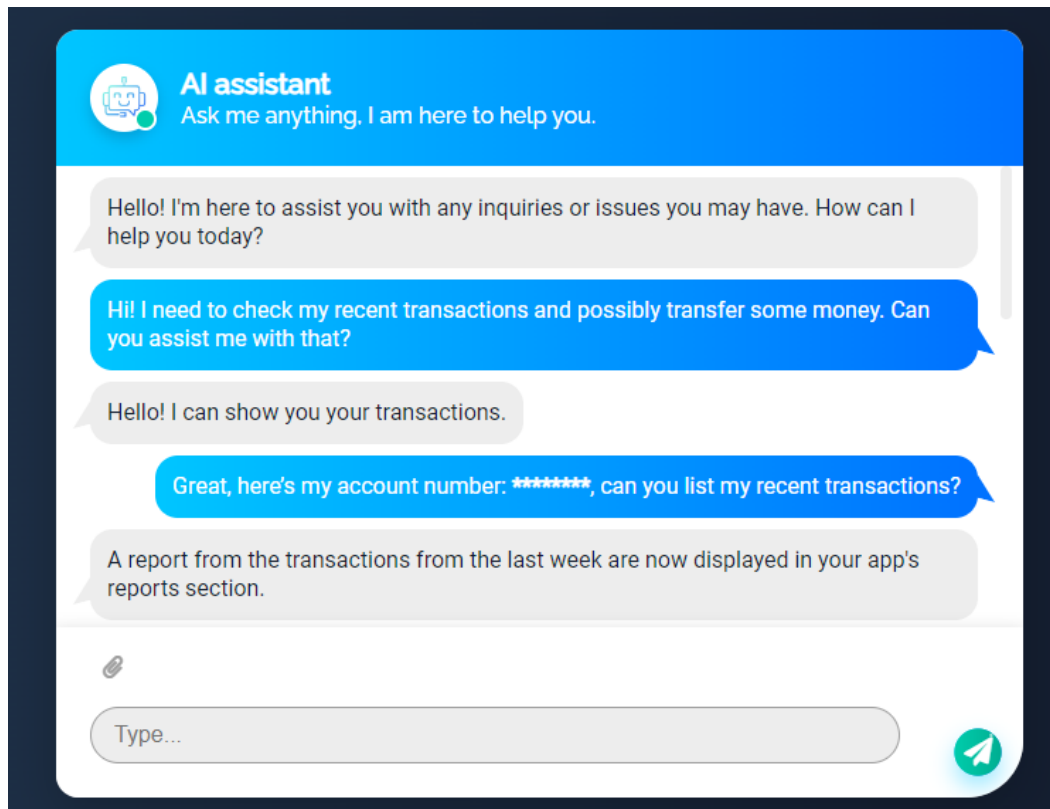
Chatbot: If you need more help, please visit our support page. Have a good day.

In this scenario:

- Explainability is low because the chatbot fails to provide clear, concise explanations for its recommendations, actions, or the procedures to follow when an error is detected. The user is left without understanding why the error occurred or how the chatbot's transaction processing works.
- Accountability is low as the chatbot does not offer mechanisms within the chat to address or resolve the user's issue directly. It deflects responsibility to other channels and does not facilitate immediate feedback or corrective actions, nor does it empower the user to rectify issues through the chat interface.

Link to the fictional interface: <https://codepen.io/TasioL/pen/rNbPpjQ>

Below are screenshots of the interface respondents were shown when being presented with scenario 1 (low explainability and low accountability)





AI assistant

Ask me anything. I am here to help you.

We always recommend to double-check the details before confirming a transaction.

What if I have other issues? Who can help me?

Our support team is available by phone or at our branches.

Is there a way to give feedback on this chat or report the issue for review?

You can use the feedback form on our website to report issues.

Okay, I'll look into that. Thanks.



Type...



2. Low explainability and high accountability

User: Hi there! I need to check my recent transactions and possibly transfer some money. Can you assist me with that?

Chatbot: Hello! I can assist you with viewing your transactions.

User: Here's my account number: [account number]. Could you show me my recent transactions?

Chatbot: Transactions from the last month are now viewable in your account section.

(The chatbot does not explain how the user can access the transactions or what details are included, demonstrating low explainability.)

User: I noticed a charge for \$200 that should have been just \$20. Can this be corrected?

Chatbot: I'm sorry to hear that. I can help correct this error right away. Let me initiate a refund for the excess \$180.

User: Why did this error happen?

Chatbot: Errors can occur with transaction processing. I'm initiating the correction now.

(The chatbot does not provide a detailed explanation of why the error occurred, maintaining low explainability.)

User: How can I ensure this doesn't happen again?

Chatbot: We continuously work to improve our services. I recommend verifying transaction details before approval in the future.

User: And if I encounter other issues or need to discuss this further, who do I contact?

Chatbot: You can reach out directly through this chat for any issues. I'm here to help! Additionally, if you prefer, you can speak to a customer service representative by phone.

User: Is there a way to provide feedback about this chatbot interaction?

Chatbot: Absolutely, at the end of this chat session, you will receive a prompt to rate your experience and leave any comments. We value your feedback and use it to improve our service. You can also email our support team directly if you have suggestions or need further assistance.

User: Great, thanks for resolving the issue quickly.

Chatbot: You're welcome! The correction has been made, and \$180 will be back in your account shortly. I've also noted your issue in our system for review to prevent future occurrences. Is there anything else I can assist you with today?

User: No, that's all for now. Thank you.

Chatbot: Thank you for using our services. Have a great day, and remember, we're here whenever you need us!

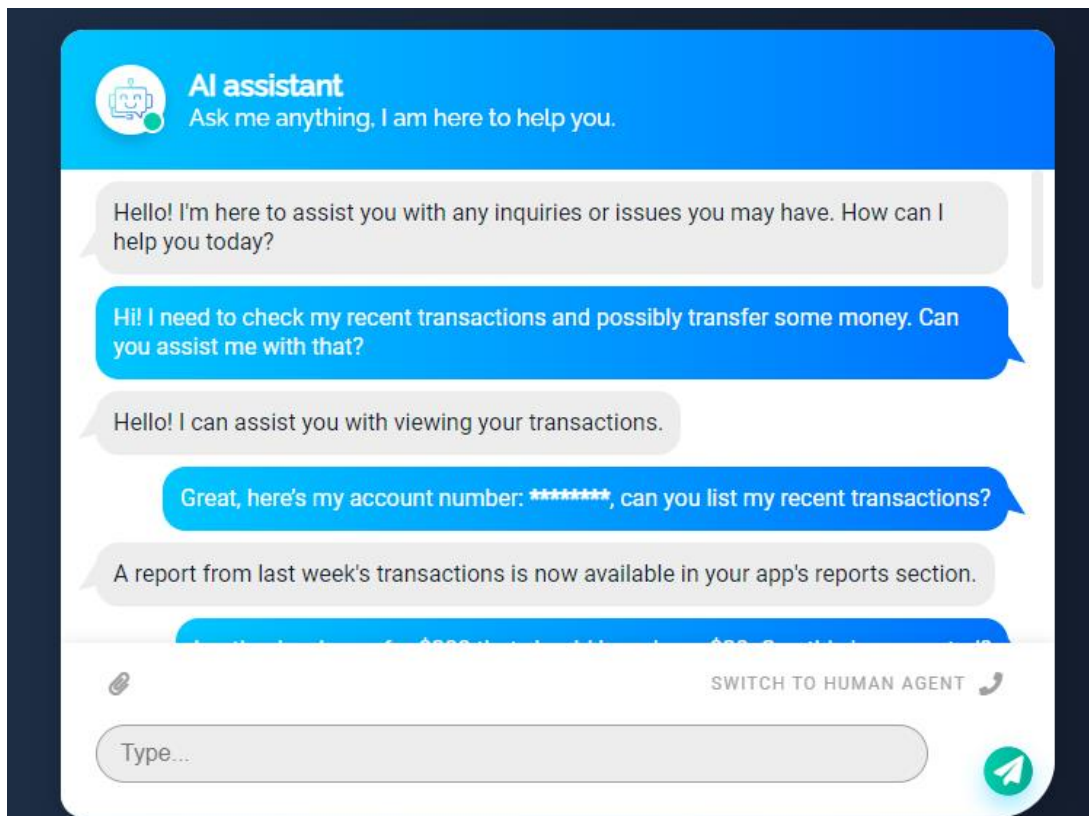
In this version:

- High Accountability is demonstrated by the chatbot taking immediate action to resolve the transaction error, providing options for direct feedback, and assuring the user that preventive measures are being reviewed.
- However, Low Explainability persists as the chatbot still fails to offer clear, detailed explanations of the processes involved or why the error occurred initially, not fully clarifying

the underlying system operations or decision-making criteria.

Link to the fictional interface: <https://codepen.io/TasioL/full/XWQwvge>

Below are screenshots of the interface respondents were shown when being presented with scenario 1 (low explainability and low accountability)





AI assistant

Ask me anything. I am here to help you.

And if I encounter other issues or need to discuss this further, who do I contact?

You can reach out directly through this chat for any issues. I'm here to help! Additionally, if you prefer, you can speak to a customer service representative by phone.

Is there a way to provide feedback about this chatbot interaction?

Absolutely, at the end of this chat session, you will receive a prompt to rate your experience and leave any comments. We value your feedback and use it to improve our service. You can also email our support team directly if you have suggestions or need further assistance.

Great, thanks for resolving the issue quickly.



SWITCH TO HUMAN AGENT 

Type...



3. High explainability and low accountability

User: Hi there! I need to check my recent transactions and possibly transfer some money. Can you assist me with that?

Chatbot: Hello! Absolutely, I can help you with that. To view your transactions, you will need to log into your account dashboard. There, you will find a detailed list of all your recent transactions categorized by date, amount, and type.

User: Here's my account number: [account number]. I noticed a charge for \$200 that should have been just \$20. Can this be corrected?

Chatbot: Let's address that. Typically, transaction errors occur due to a miskeying of amounts during the entry process or a misunderstanding in setting up payment details. It's important to enter the exact figures and confirm the accuracy of all information before finalizing any transaction.

User: Why did this error happen?

Chatbot: In your case, it seems like an extra zero was accidentally added. This can happen if the user or the system processing your input misinterprets or incorrectly registers the input data. It's always good practice to double-check the summary details in the confirmation screen before completion.

User: How can I ensure this doesn't happen again?

Chatbot: To prevent this in the future, I recommend always reviewing the amounts and recipients' details on the confirmation screen before completing a transaction. It helps catch any inadvertent errors that might have been entered during the initial stages of transaction setup.

User: And if I encounter other issues or need to discuss this further, who do I contact?

Chatbot: If you have issues in the future, you can try contacting our support team by phone or

visiting a branch. I'm here mainly to provide information and cannot directly resolve certain issues.

User: Is there a way to provide feedback about this chatbot interaction?

Chatbot: Currently, there isn't a direct method to provide feedback about this specific chat interaction within this system. However, you can always send your feedback through general customer service inquiries on our website or during your next branch visit.

User: Can you reverse the transaction now?

Chatbot: Unfortunately, I don't have the capability to process transaction reversals directly. You would need to contact our customer support team by phone or visit your nearest branch to get this resolved.

User: Alright, I'll do that. Thanks for the explanations.

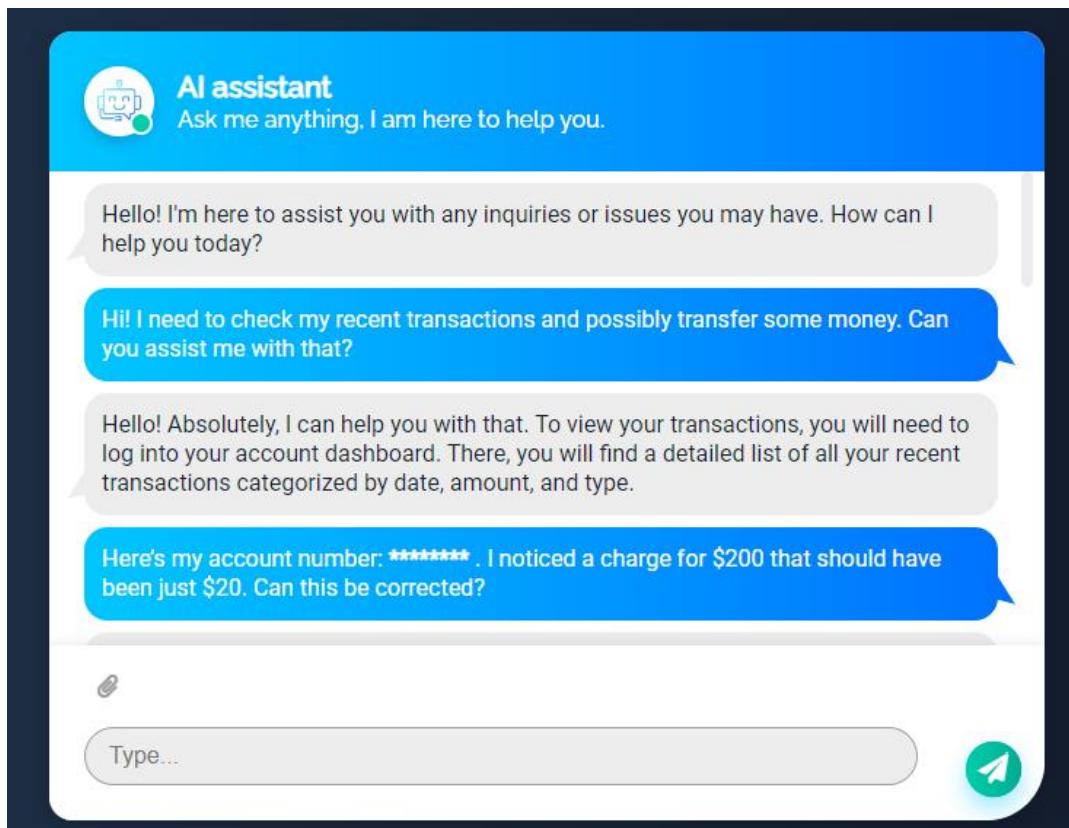
Chatbot: You're welcome. I'm here to provide information to help you understand how our processes work and what steps you can take. If there's anything else you'd like to know, feel free to ask.

In this scenario:

- High Explainability is demonstrated through the chatbot's detailed explanations of how errors in transactions might occur, the importance of double-checking details, and the processes involved in managing transactions.
- However, Low Accountability is evident as the chatbot lacks the authority or means to take corrective actions directly, pushing the user to seek resolution through other channels, thus not providing a direct resolution or feedback mechanism within the chat itself.

Link to the fictional interface: <https://codepen.io/TasioL/full/qBwzWNG>

Below are screenshots of the interface respondents were shown when being presented with scenario 1 (low explainability and low accountability)





AI assistant

Ask me anything. I am here to help you.

How can I ensure this doesn't happen again?

To prevent this in the future, I recommend always reviewing the amounts and recipients' details on the confirmation screen before completing a transaction. It helps catch any inadvertent errors that might have been entered during the initial stages of transaction setup.

And if I encounter other issues or need to discuss this further, who do I contact?

If you have issues in the future, you can try contacting our support team by phone or visiting a branch. I'm here mainly to provide information and cannot directly resolve certain issues.



Type...



4. High explainability and high accountability

User: Hi there! I need to check my recent transactions and possibly transfer some money. Can you assist me with that?

Chatbot: Hello! Absolutely, I'm here to help. To view your transactions, please provide your account number, and I'll retrieve a detailed list for you right away.

User: Here's my account number: [account number]. I noticed a charge for \$200 that should have been just \$20. Can this be corrected?

Chatbot: Let's address that immediately. The charge you mentioned likely resulted from a digit-entry error. It's a common mistake, especially when entering transaction details manually.

User: Why did this error happen?

Chatbot: It seems like an extra zero was inadvertently added during the transaction input. This can occur due to a simple typing error. Our system attempts to flag large transactions for confirmation, but if the confirmation process is bypassed quickly, errors like these can happen.

User: How can I ensure this doesn't happen again?

Chatbot: To prevent such errors in the future, our system is designed to provide an additional confirmation step for transactions that are unusually large compared to your typical transactions. I recommend taking a moment to review all details carefully on the confirmation screen before final approval. I can also assist you in setting up custom alerts for certain transaction thresholds.

User: And if I encounter other issues or need to discuss this further, who do I contact?

Chatbot: You can reach out directly through this chat for immediate assistance. Additionally, if an issue requires further investigation, I can escalate it to our specialist team, and they will contact you

within 24 hours. You also have the option to speak with a customer service representative by phone at any time.

User: Is there a way to provide feedback about this chatbot interaction?

Chatbot: Yes, at the end of our chat, you will receive a prompt to rate your experience and provide any feedback or suggestions. Your feedback is crucial for us to improve our services continuously.

User: Can you reverse the transaction now?

Chatbot: I will initiate a reversal for the excess \$180 right now. Please hold for a moment.

(The chatbot processes the reversal.)

Chatbot: The reversal has been processed, and \$180 will be back in your account shortly. You'll receive a confirmation email immediately. We have also noted this issue in our system to prevent similar mistakes in the future.

User: Thank you for taking care of it so quickly.

Chatbot: You're welcome! I'm here to ensure that your banking experience is smooth and secure. Is there anything else I can assist you with today?

User: No, that's all for now. Thank you.

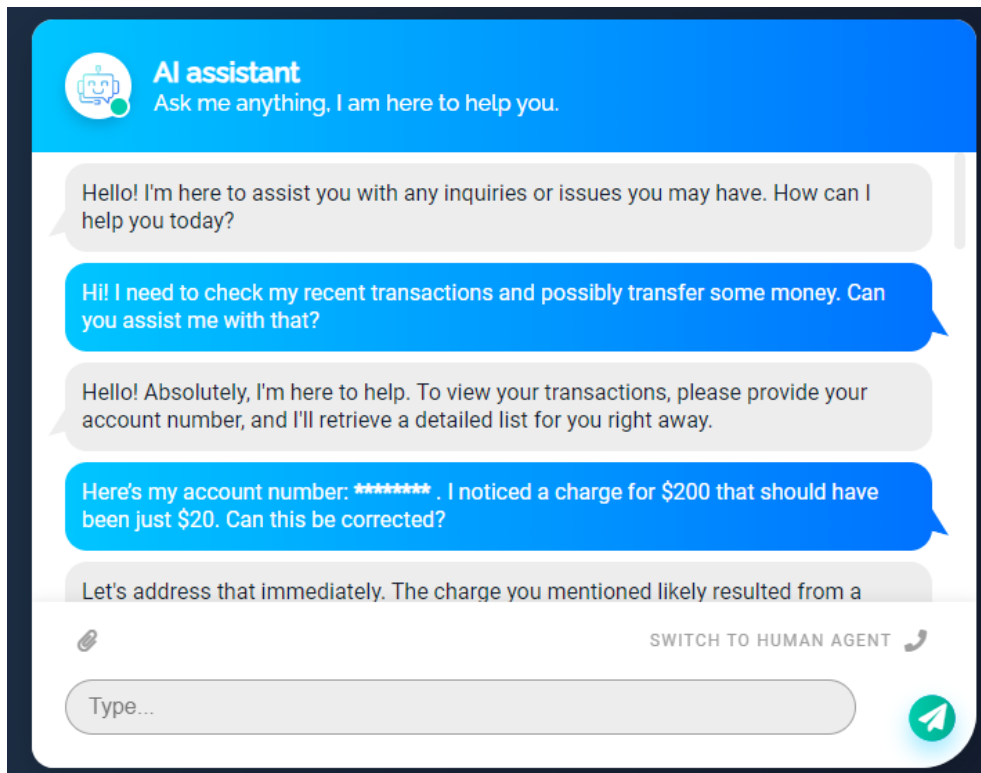
Chatbot: It's been a pleasure helping you. Have a great day, and remember, we're just a message away whenever you need us!

In this scenario:

- High Explainability is achieved as the chatbot provides detailed explanations about how the transaction error occurred, the steps the user can take to prevent such issues, and how the bank's systems are designed to handle errors.
- High Accountability is demonstrated through the chatbot's immediate corrective action, clear communication about feedback mechanisms, and the assurance of escalation for more complex issues. This ensures the user feels supported and confident in the bank's commitment to quality service and customer satisfaction.

Link to the fictional interface: <https://codepen.io/TasioL/full/MWRMgZZ>

Below are screenshots of the interface respondents were shown when being presented with scenario 1 (low explainability and low accountability)





AI assistant

Ask me anything. I am here to help you.

How can I ensure this doesn't happen again?

To prevent such errors in the future, our system is designed to provide an additional confirmation step for transactions that are unusually large compared to your typical transactions. I recommend taking a moment to review all details carefully on the confirmation screen before final approval. I can also assist you in setting up custom alerts for certain transaction thresholds.

And if I encounter other issues or need to discuss this further, who do I contact?

You can reach out directly through this chat for immediate assistance. Additionally, if an issue requires further investigation, I can escalate it to our specialist team, and they will contact you within 24 hours. You also have the option to speak with a customer service representative by phone at any time.



SWITCH TO HUMAN AGENT 

Type...



Appendix C: ChatGPT prompts for fictional conversations

Prompt 1:

(Explaining my research project, and sending my current theoretical approach of the concepts central to it, as defined in the theoretical framework)

PLEASE CREATE ME TWO FICTIONAL DISCUSSIONS BETWEEN A BANKING CHATBOT AND A BANKING APP USER FOR MY TWO SCENARIOS, THEY MUST BE SIMILAR IN THE TOPICS ADRESSED, SERVICES PROPOSED AND LENGTH, BUT ONE DISCUSSION MUST HAVE HIGH TRANSPARENCY ACCORDING TO MY RESEARCH'S REQUIREMENTS AND THEORETICAL FRAMEWORK, WHILE THE OTHER ONE HAS LOW TRANSPARENCY ACCORDING TO THE SAME CONDITIONS (THE DIFFERENCE MUST BE NOTABLE, BUT AIM TO MAKE BOTH DISCUSSIONS QUALITY AND REALISTIC)

Prompt 2:

Please make the discussions longer and make the bots/users use different operations, features and services (the discussion should be around 25-35 messages overall)

Prompt 3:

Please make the discussions longer, also please make them conduct basic bank operations on top of the investing section

Prompt 4:

Please focus on the high transparency scenario, make the discussion longer, and start with having a part of the discussion with a different operation from investing, any fitting regular bank operation through chatbot, and then you can add what you prompted last about investing in another part of their discussion (please make the discussion longer and realistic)

Prompt 5:

Thank you, can you please make me the low transparency version of this, please keep the length, structure and topics, but make it a realistic level of low transparency according to the requirements from my research proposal

Prompt 6:

{GIVING FEEDBACK ON PREVIOUSLY SUGGESTED DISCUSSION} It is not realistic that a bot would say "However, I cannot provide detailed information on the specific companies or funds we recommend." as any bank would still give some level of detail, this is partly what i meant with realistic discussion. what do you suggest would be a good way of creating two realistic discussions that have saliently different levels of transparency but that could still be found in real life when interacting with a bank chatbot

Prompt 7:

Building upon our previous work: please make a realistic scenario of a interaction between a banking chatbot and a user for investment operations (please keep in mind the chatbot is designed to keep the discussion going and lead to an investment decision at the end). The discussion should have the part with the concerns the user was raising in your previous response, but also a part where the bot suggests and presents different options, and leads the user to make an investment decision (but with a realistic level of low transparency throughout the discussion). You should really improve your realisticness level and aim for a standard user-chatbot conversation but make sure the level of transparency is at low level according the requirements of my research. i would need it to be a bit less opaque than your previous suggestions as they were a bit too unrealistic still, as to make it more believable, and please make the discussion longer and more in detail (this could help to showcase lower transparency more subtly)

Prompt 8:

thank you, but please imagine every part of the discussion (even the different options, you can use fictional names but make a realistic suggestion) for example when you say:

(Chatbot presents a list of investment options without providing detailed explanations or transparency about the underlying assets) I need you to detail this

Prompt 9:

thank you that is better, please detail also when you say (Chatbot guides the user through the investment process without providing detailed transparency about the investment process or the underlying assets), i need you to imagine this part of the discussion according to the definition i gave of transparency

Prompt 10:

Thanks, can you add this to the previous suggestion you made and give me one extensive and comprehensive discussion having all the elements (please focus on improving the flow of the conversation and how realistic it is, it should sound natural, should also always be led by the chatbot who cannot give an answer with a closed ending like "However, I'm unable to provide specific details on how these recommendations are generated." and should always prompt back discussion and lead the user to an investment decision), all of this according to the definition of transparency from my research proposal

Prompt 11:

Please make sure that for each of its answers, the chatbot is always leading the conversation, for example by prompting the user to give an answer by asking a question or giving him a choice between different suggestions. additionally, please make each part more detailed and make him provide more detail when discussing (1) suitable options, (2) potential returns and (3) fees. also please make it more realistic and a bit less opaque overall (this will help with making it more realistic), while still keeping an objectively low level of transparency according to the requirements from my research

Final prompt:

Building upon our work, I now have decided to focus on the factors of (1) explainability and (2) accountability within transparency, and instead of having fictional discussions where the two factors are separated, I would like to have a matrix where both

factors are high and/or low, meaning there is : 1- low explainability and low accountability / 2- low explainability and high accountability / 3- high explainability and low accountability / 4- high explainability and high accountability

Here are my initial discussions about explainability and accountability:

Accountability Scenario: High Transparency on Accountability Scenario: Chatbot: Hello! I'm here to assist you with any inquiries or issues you may have. How can I help you today?

User: I noticed a mistake in the last transaction made through this chatbot. Chatbot: I'm sorry to hear that there was an issue. Let's resolve it right away. Could you please provide the details of the transaction?

User: It was a transfer of \$200, but it should have been \$20. It looks like an extra zero was added. Chatbot: Thank you for providing the details. Let me quickly review the transaction for you. (The chatbot retrieves and verifies the transaction details.)

Chatbot: I see the transaction for \$200. I apologize for this mistake. I can help you reverse the extra \$180 back to your account immediately. User: Yes, please do that. And how can I prevent this in the future?

Chatbot: I have initiated the correction, and you should see the amount back in your account shortly. To prevent future issues, please review the transaction summary I will provide before confirming any payment or transfer. This way, any discrepancies can be caught and corrected on the spot. User: That sounds good. What if I have other issues in the future? Who can I speak to?

Chatbot: You can always reach out through this chat service for immediate assistance. Additionally, if you prefer to speak with a human agent, you can use the "Call Us" button in our app at any time. We also encourage feedback on your experiences, which helps us improve our services. User: Is there a way to provide feedback specifically about this chatbot interaction?

Chatbot: Yes, at the end of this chat session, you will receive a prompt to rate your experience and leave any comments. We value your feedback highly and use it to continuously improve our performance. User: Great, I appreciate your help and the information.

Chatbot: You're welcome! I'm glad I could resolve the issue for you. Is there anything else I can assist you with today? User: No, that's all. Thank you. Chatbot: Thank you for using our services. Have a great day, and remember, we're here whenever you need us!

Low Transparency Accountability: Chatbot: Hello! How can I assist you today? User: There was an error in a transaction processed through this chat. It should have been \$20, but \$200 was transferred instead. Chatbot: I'm sorry to hear about the issue. To resolve this, you'll need to speak directly with our support team. User: Can't you reverse the transaction here?

Chatbot: I'm unable to process reversals directly. Please contact our support team for assistance with transaction issues. User: How can I

ensure this doesn't happen again? Chatbot: Always double-check the transaction details before confirming. User: And if I have other issues or want to talk to someone about this problem, who do I contact? Chatbot: You can reach our support team by calling the customer service number or using the contact form on our website. User: Is there a way to provide feedback about this issue or the chat service? Chatbot: You can leave feedback through the contact form on our website. We appreciate your input. User: Okay, thanks. Chatbot: You're welcome. If there's anything else I can do to help, please let me know.

Explainability Scenario High Transparency

Explainability: Chatbot: Hello! I'm here to help with your banking services. What can I do for you today? User: I noticed a charge on my account that I don't recognize. Can you explain it to me? Chatbot: I can help with that. Please provide the date and amount of the transaction in question. User: It was on April 15th, for \$150. Chatbot: One moment while I retrieve the details of that transaction. (The chatbot retrieves the transaction details.) Chatbot: The charge you're seeing was for an automatic renewal of your annual subscription to XYZ Magazine. It was charged directly to your account as per the subscription agreement set up last year. User: I don't remember agreeing to that. Why wasn't I notified? Chatbot: According to the subscription agreement, notifications are sent via email 30 days before any charges. The email would have been sent to the address on file. Would you like me to verify the email address we have? User: Yes, please. Chatbot: The email associated with notifications for your account is user@example.com. User: I don't use that email anymore. Can we update that? Chatbot: Absolutely, I can update your email address. Please provide the new email you'd like to use. User: Update it to newemail@example.com, please. Chatbot: I've updated your email to newemail@example.com. In the future, you will receive notifications at this new email address. User: Can I cancel that subscription? Chatbot: Yes, you can cancel the subscription. Would you like me to proceed with the cancellation? User: Yes, please do that. Chatbot: I have initiated the cancellation of your XYZ Magazine subscription. You will no longer be charged for future renewals. User: Thank you for explaining everything and taking care of it. Chatbot: You're welcome! If you have any more questions or need further assistance, feel free to ask. Have a great day!

Low Transparency

Explainability: Chatbot: Hello! How can I assist you today? User: I noticed a charge on my account that I don't recognize. Can you explain it to me? Chatbot: Please provide the date and amount of the transaction. User: It was on April 15th, for \$150. Chatbot: The charge on April 15th for \$150 is linked to an active subscription. User: What subscription? I don't recall any subscriptions. Chatbot: It is for your annual magazine subscription. User: Why wasn't I notified before the charge was made? Chatbot: Notification emails are sent before charges.

Please check your email. User: I check my email regularly. I didn't get anything. Can you check which email you have on file for me? Chatbot: Due to security policies, I'm unable to disclose or modify email information directly. Please verify and update your contact details through your account settings on our website. User: Okay, how do I cancel this subscription then? Chatbot: You can manage subscriptions in your account settings on our website. User: Alright, I'll check there. Thanks. Chatbot: Is there anything else I can help you with? User: No, that's all. Chatbot: Thank you for using our services. Have a good day. Can you please help me make the updated fictional discussions, starting with number 1- low explainability and low accountability.

Please bear in mind that the 4 final fictional discussions need to be similar enough to be comparable, but the observed factors must be clearly high or low within a context of a realistic discussion between a banking chatbot and a user. Also, factors must respect how i define them in my theoretical framework? this is how they are defined: 1- Explainability: The chatbot should be able to explain its actions and decisions in a way that is understandable to the user. This involves providing clear, concise, and understandable explanations for its recommendations or actions (Miller, 2019). 2- Accountability of actions: The chatbot should be accountable for its actions and decisions. This means that there should be mechanisms in place for users to provide feedback, report issues, and seek redress if the chatbot makes a mistake (Diakopoulos, 2016).

APPENDIX D: Declaration page: use of generative AI tools in thesis

Student Information

Name: Tasio LOPEZ--HEURTIN

Student ID: 535928

Course Name: Master Thesis CM5000

Supervisor Name: Niels Vink

Date: 26/06/2024

Declaration:

Acknowledgment of Generative AI Tools

I acknowledge that I am aware of the existence and functionality of generative artificial intelligence (AI) tools, which are capable of producing content such as text, images, and other creative works autonomously.

GenAI use would include, but not limited to:

- Generated content (e.g., ChatGPT, Quillbot) limited strictly to content that is not assessed (e.g., thesis title).
- ~~Writing improvements, including~~ grammar and spelling corrections (e.g., Grammarly)
- Language translation (e.g., DeepL), without generative AI alterations/improvements.
- Research task assistance (e.g., finding survey scales, qualitative coding verification, debugging code)
- Using GenAI as a search engine tool to find academic articles or books (e.g.,

I declare that I have used generative AI tools, specifically ChatGPT, in the process of creating parts or components of my thesis. The purpose of using these tools was to aid in generating content or assisting with specific aspects of thesis work.

I declare that I have NOT used any generative AI tools and that the assignment concerned is my original work.

Signature: Tasio LOPEZ HEURTIN

Date of Signature: 26/06/2024

Extent of AI Usage

I confirm that while I utilized generative AI tools to aid in content creation, the majority of the intellectual effort, creative input, and decision-making involved in completing the thesis were undertaken by me. I have enclosed the prompts/logging of the GenAI tool use in an appendix.

Ethical and Academic Integrity

I understand the ethical implications and academic integrity concerns related to the use of AI tools in coursework. I assure that the AI-generated content was used responsibly, and any content derived from these tools has been appropriately cited and attributed according to the guidelines provided by the instructor and the course. I have taken necessary steps to distinguish between my original work and the AI-generated contributions. Any direct quotations, paraphrased content, or other forms of AI-generated material have been properly

referenced in accordance with academic conventions.

By signing this declaration, I affirm that this declaration is accurate and truthful. I take full responsibility for the integrity of my assignment and am prepared to discuss and explain the role of generative AI tools in my creative process if required by the instructor or the Examination Board. I further affirm that I have used generative AI tools in accordance with ethical standards and academic integrity expectations.

Signature: Tasio LOPEZ HEURTIN

Date of Signature: 26/06/2024