[Inner endpaper]

HEC MONTRÉAL

Interaction Friction in AI: Improving User Experience and Decision Making

par

Pauline Françoise Marie Malaguti

Pierre-Marjorique Leger
HEC Montréal
Co-Directeur de recherche

Sylvain Sénécal
HEC Montréal
Codirecteur de recherche

Sciences de la gestion
Data Science for Business and Business Analytics

Mémoire présenté en vue de l'obtention
du grade de maîtrise ès sciences en gestion
(M. Sc.)

August 2024

© Pauline Francoise Marie Malaguti, 2024

Résumé

Ce mémoire explore le concept de « friction d'interaction » dans les interactions humain-AI, avec pour objectif d'améliorer la prise de décision des utilisateurs. Il est structuré en deux articles principaux.

Le premier article analyse l'impact de la présence ou de l'absence de l'AI ainsi que des différents niveaux de friction d'interaction sur les perceptions des utilisateurs, notamment en termes d'utilité, de confiance, de sentiment de contrôle, de performance, et de charge cognitive. À travers une étude expérimentale impliquant des agents conversationnels (chatbots) AI textuels, les résultats montrent que la friction contrôlée améliore les perceptions des utilisateurs du système ainsi que les processus de prise de décision. Ces résultats suggèrent que l'ajout délibéré de friction peut contrebalancer les effets négatifs potentiels des interactions AI trop fluides, tels que la réduction de l'autonomie des utilisateurs et la complaisance induite par l'AI.

Le second article porte sur l'amélioration des interactions humain-ordinateur (HCI) en milieu de travail. Il évalue comment les systèmes AI peuvent être conçus non seulement pour permettre aux employés d'accomplir efficacement leurs tâches, mais aussi pour les soutenir dans leurs processus de prise de décision. L'étude met en avant l'importance de l'équilibre entre l'efficacité et les principes de conception centrés sur l'humain pour favoriser un sentiment de contrôle et d'appropriation chez les utilisateurs.

Dans l'ensemble, ce mémoire contribue au domaine de l'HCI en démontrant que l'intégration de friction d'interaction contrôlée dans les systèmes AI peut améliorer l'expérience utilisateur et les performances. Les résultats offrent des insights précieux pour la conception de technologies AI à la fois efficaces et habilitantes, encourageant des interactions plus réfléchies et engagées entre les humains et l'AI.

Mots-clés : Conception de friction AI, Interactions humain-AI, État cognitif, Utilité, Confiance, Sentiment de contrôle, Performance, Assistant d'Intelligence Artificielle Générative (GENAIA).

Méthodes de recherche : Expérience en laboratoire ; pupillométrie ; questionnaires.

Abstract

This thesis explores the concept of 'interaction friction' in AI-human interactions, aiming to improve users' decision-making. The paper is structured into two main articles.

The first article investigates how the presence or absence of AI as well as varying levels of interaction friction affect user perceptions of usefulness, confidence, trust, sense of agency, performance and cognitive load. Through an experimental design involving AI text-based chatbots, the study finds that controlled friction improves user's perceptions of the system and decision-making processes. The results suggest that deliberate friction can counteract the potential negative effects of overly smooth AI interactions, such as reduced user agency and AI-led complacency. The second article focuses on enhancing Human-Computer Interaction (HCI) in the workplace. It evaluates how AI systems can be designed to not only help employees to perform tasks efficiently but also support them in their decision-making processes. The study emphasises the importance of balancing efficiency with human-centric design principles to foster a sense of control and ownership among users.

Overall, this thesis contributes to the field of HCI by demonstrating that incorporating controlled interaction friction in AI systems can lead to better user experience and performance. The findings provide valuable insights for designing AI technologies that are both efficient and empowering, promoting more deliberate and engaged interactions between humans and AI.

Keywords : AI Design Friction, AI-Human Interactions, Cognitive State, Usefulness, Confidence, Trust, Sense of Agency, Performance, Generative Artificial Intelligence Assistant (GENAIA).

Research methods : Lab Experiment ; Pupillometry ; Questionnaires

Table of contents

List of tables and figures

List of abbreviations and acronyms

Artificial Intelligence = AI

Confidence in AI = CAI

Generative AI Assistant = GENAIA

Human-Centred AI = HCAI

Human-Computer Interaction = HCI

Perceived Usefulness = PUS

Sense of Agency = SOA

Trust in AI = TAI

User Experience = UX

Preface

The HEC Montréal Research Ethics Board (REB) approved this project (Certificate # 2023-5385) and authorised to conduct the experiment presented in this thesis in May 2023.

The Administrative Director of the Master of Science in Data Science for Business and business analytics authorised the following dissertation to be written in article form. The articles aim to understand the impact of friction in AI interaction on users' perceptions, cognitive state and task performance.

Acknowledgements

This thesis marks the culmination of several years of work and study, during which I have experienced immense growth as a student, researcher, and individual. I am deeply grateful to all those who have supported me along this journey.

First and foremost, I would like to extend my sincere gratitude to my supervisors, Pierre-Majorique Léger and Sylvain Sénécal. Their expertise, guidance, and encouragement have been invaluable throughout this project. Their enthusiasm and insightful feedback have consistently pushed me to improve and achieve more than I thought possible.

A special thank you goes to Alexander J. Karran, whose support and guidance were crucial in shaping the direction of this project. Your assistance and insightful discussions have been a keystone of my research, and I am deeply appreciative of your contributions.

I am also profoundly grateful to my friends for their constant support and presence. Elise, thank you for being there through the highs and lows, and for your encouragement. Your friendship has been a source of strength for me.

To my fellow lab students and friends, thank you for sharing this journey with me. Together, we have formed a strong team that I could rely on for motivation and encouragement. Your collective strength and determination have inspired me to persevere, and I am thankful for the solidarity we have built.

To my family, I owe the deepest thanks. Your unconditional love and support have been my foundation. Thank you for believing in me and for always being there to cheer me on.

Lastly, I would like to acknowledge the financial support provided by NSERC, PROMPT, and other industrial partners. Their contributions have been instrumental in the completion of this research.

Introduction

"Can machines think?" This important question posed by Alan Turing in his seminal 1950 paper laid the foundation for the field of artificial intelligence. Today, Artificial Intelligence (AI) has become an integral part of everyday life, transforming various sectors from healthcare to finance, education, and beyond. In healthcare, AI-powered diagnostic tools assist doctors in making more accurate diagnoses and robotic surgeries are becoming more common (Fujita, 2020). In finance, AI algorithms manage investments, detect fraudulent transactions, and offer personalised financial advice (Goodell et al., 2021). Education has also seen a significant impact, with AI-driven tutoring systems providing customised learning experiences, grading automation reducing the workload for educators, and AI tools facilitating online learning platforms (Zhang & Aslan, 2021).

The interaction between humans and AI systems, known as Human-AI Interaction (HCAI), has traditionally focused on enhancing efficiency and functionality, following the ‘Don’t Make Me Think’ paradigm (Krug, 2005). These advancements are designed to create seamless and intuitive user experiences that minimise effort and maximise productivity (Messeri et al., 2021). However, as AI becomes more integrated in daily activities, there is a growing need to understand the nuanced impacts of these interactions on human perceptions, emotions, and decision-making processes.

As AI systems take on more roles traditionally performed by humans, the balance between human control and machine autonomy becomes increasingly fragile (Roto et al., 2019). There is a critical need to explore how AI can be designed to not only perform tasks efficiently but also to support and enhance human decision-making and emotional well-being (Ben Shneiderman, 2022). This requires a shift in perspective from the traditional focus on frictionless interactions to an approach that considers the benefits of deliberate interaction friction. In physics, friction is defined as “the force that works against an object as it slides along the surface of another object” (Cambridge University Press, 2024). This concept can be extended to the realm of human-computer interaction. Interaction friction, in the field of human AI interaction, refers to the intentional inclusion of elements that require user input or engagement, encouraging a more active and thoughtful interaction (Cox et al., 2016). By incorporating controlled friction, AI systems can prompt users to engage more deeply with the tasks at hand, encouraging critical thinking and enhancing their overall decision-making process (Naiseh et al., 2021).

This thesis explores the concept of 'interaction friction' in AI-human interactions, aiming to enhance user engagement and decision-making. By introducing controlled friction, such as prompts for user feedback or decision points, AI systems can encourage users to engage more deeply with the tasks at hand, potentially leading to better outcomes (Friedland, 2020). This research seeks to identify how varying levels of interaction friction affect user perceptions of usefulness, confidence, trust, sense of agency, and cognitive load. The goal is to provide insights that can guide the design of AI systems that are not only efficient but also empowering and supportive of human users in their decision-making processes.

Traditional HCI research has aimed to create frictionless interactions to maximise efficiency. The goal has been to streamline user experiences, making interactions with technology as smooth and effortless as possible. This approach assumes that reducing user effort leads to greater satisfaction and productivity. However, this focus on frictionless design can lead to unintended consequences, such as diminished user agency and responsibility. When interactions are overly simplified, users may become passive recipients of information rather than active participants in the decision-making process. This can result in a decreased sense of control and ownership over the outcomes of their actions (Friedland, 2020). Moreover, research into neurocognitive activity has shown that some users' cognitive engagement can become so reduced that they rely solely on AI output and suggestions for decision making, leading to phenomena known as AI-led complacency and automation bias (Pattyn et al., 2008). The consequences of this complacency have already been observed, as highlighted by a Forbes report where 43% of the companies surveyed expressed concerns about their increasing reliance on AI for their daily operations (Haan, 2023).

As AI systems become more prevalent, their role in shaping human decisions and behaviours gains critical importance. Educational AI systems, for instance, offer personalised learning experiences but can also influence what and how students learn, potentially impacting their critical thinking and critical thinking skills (Holmes & Tuomi, 2022). Similarly, AI-driven financial tools can facilitate investment decisions and financial planning, yet they may also lead users to overly rely on automated suggestions, reducing their engagement and understanding of financial matters (Buckley et al., 2021).

This research builds on foundational concepts in HCI and AI, including Human-Centred AI (HCAI). The goal of HCAI has been defined as "Human-Centered AI (HCAI) focuses on understanding purposes, human values and desired AI properties in the creation of AI systems by applying Human-Centered Design practices. HCAI seeks to augment human capabilities while maintaining human control over AI systems, by considering the necessity, context, and ethical and legal conditions of the AI system as well as promoting individual and societal well-being" (Schmager et al., 2023). HCAI emphasises designing AI systems that enhance human capabilities, ensuring they are reliable, understandable, and empowering for users (B. Shneiderman, 2022). By integrating human factors and ethical design principles, HCAI aims to create AI technologies that are not only effective but also aligned with human values and needs (Amershi et al., 2019). This approach underscores the importance of transparency, accountability, and user autonomy in the development of AI systems.

The evolution of chatbots and AI-driven decision support systems provides a contextual background for this study, highlighting the need for AI interfaces that balance efficiency with human-centric design principles. Early chatbots like ELIZA and PARRY were designed primarily for simple, text-based interactions, evolving into more sophisticated systems such as SmartChild and voice assistants, which offer a wide range of functionalities including task management and information retrieval (Adamopoulou & Moussiades, 2020). These advancements demonstrate the potential for AI to assist in complex decision-making processes, yet they also reveal the limitations of frictionless design.

The concept of 'interaction friction' is introduced, challenging the prevailing notion that minimising user effort maximises efficiency. Traditional HCI approaches have followed the "Don't make me think" approach, that encourages seamless and effortless interactions (Krug, 2005). It follows the assumption that reducing cognitive load leads to higher user satisfaction and productivity (Sun et al., 2022). Instead, this study posits that deliberate friction can enhance cognitive engagement and decision-making clarity. By introducing controlled elements of friction, such as prompts for user input or decision points, AI systems can encourage users to process information more thoroughly and reflect on their choices. This approach aligns with theories of cognitive load and human-computer interaction (HCI), which suggest that a certain level of engagement and challenge can enhance learning and performance. For example, by requiring users to verify information or make active decisions, AI systems can foster a sense of responsibility and ownership over the outcomes, leading to improved user satisfaction and trust.

Despite progress in AI, there remains a significant gap in understanding how deliberate interaction friction can enhance user engagement, cognitive clarity, and decision-making quality (Natali, 2023). Interaction friction, when applied thoughtfully, can encourage users to engage more deeply with the content and decisions they are making. For example, requiring users to verify information or make choices at key decision points can foster a sense of involvement and responsibility (Mejtoft et al., 2019). This deliberate design choice can help users retain a greater sense of agency and make more informed decisions.

Understanding the balance between frictionless and friction-filled interactions is crucial for designing AI systems that not only perform tasks efficiently but also support and empower users. This study aims to bridge this gap by investigating the impact of interaction friction on user engagement and decision-making. The research explores how different levels of interaction friction affect users' perceptions of usefulness, confidence, trust, and cognitive load, ultimately aiming to provide a framework for designing AI systems that enhance both efficiency and user empowerment.

To explore the impact of different AI interaction modes on user perceptions and decision-making, this thesis is guided by the following general research question:

To what extent do various AI friction affect user perceptions, cognitive state, and behaviour?

This question aims to provide a comprehensive understanding of how varying levels of interaction friction and different AI interaction modes can enhance user engagement and decision-making processes. The detailed research questions addressing specific aspects of this overarching inquiry are explored within the individual articles comprising this thesis.

The research employs a mixed-methods approach, combining quantitative, observational and physiological data to provide a comprehensive analysis of user perceptions and cognitive responses. Various AI interaction modes were designed through different designs of Generative AI Assistant (GENAIA). Participants' perceptions of usefulness, confidence, trust, sense of agency, cognitive load, and task completion time are measured. In addition, physiological measurements provide insights into the cognitive state of users, offering a nuanced understanding of the dynamic interplay between user perceptions, performance and AI interactions.

This thesis is organised into two main articles, each addressing specific angles of understanding of the impact of human-ai interaction. The first article, "Interaction Friction in AI: Improving User Experience and Decision Making," examines the effects of introducing interaction friction in AI-human interactions. It investigates how different interaction modes impact user engagement, cognitive states, and decision-making processes. The second article, " Enhancing Human-Computer Interaction in the Workplace" provides insights and examples on how the different designs of AI frictions can be used in the Workplace to improve user's experience and performance.

The table below presents the percentage contributions of the student alongside those of the other team members across all components of the research project.

Component	Distribution
Research Questions	● Industrial partner initially provided concepts of interests ● The student and the supervisors converted these concepts into research questions based on the literature
Literature Review	Research on scientific databases to find relevant articles related to the constructs and research questions 100%
Conception and Experimental Design	Application to the Research Ethics Board (REB) of HEC Montréal. 100% ● Preparation of documentation related to the submission by the student. ● Application reviewed by thesis co-supervisors and Tech3Lab operations staff. Preparation of the data collection ● Stimuli provided by research partner ● Installation and testing of the stimuli 100%

	● Creation of the questionnaires on Qualtrics 90%
Participant Recruitment	Recruiting participants for the study 20% ● The Participants were recruited through the Tech3Lab's panel ● The student recruited last minute participants to replace cancellations
Data Collection	Involvement in the Data Collection 80% ● The student was involved in the data collection as a moderator ● The student was also present to troubleshoot problems with the stimuli ● The installation of the collection tools was performed by Tech3Lab research assistants
Data Analysis	● Exporting and preprocessing of the data into usable data frames on Python 100% ● Data Analysis in R 100%
Drafting the Thesis	Writing of the introduction, scientific article, managerial article, and conclusion 100%

Table 1: Table of Contribution

Chapter 1:
Interaction Friction in AI: Improving User Experience and Decision Making

Pauline Françoise Marie Malaguti, Sylvain Sénécal, Pierre-Majorique Léger

HEC Montréal

Abstract

This study examines the effects of introducing interaction friction in AI-human interactions to enhance user perceptions and decision-making, aiming to explore how manipulating interaction modes can create a more mindful state and improve cognitive state during AI use. Using a Generative Artificial Intelligence Assistant (GENAIA), we investigated two factors: AI vs. Non-AI and Non-Friction AI (regular chatbot mode) vs. Friction AI (summary and form chatbot modes). Participants' perceptions of usefulness, confidence, trust, sense of agency, cognitive load, and performance were measured. Results showed that AI interaction modes were generally perceived as more useful than the control condition, although adding friction to chatbots did not significantly enhance perceived usefulness. Confidence and trust were higher in AI-assisted modes, but differences among the AI modes were not significant. Cognitive load was significantly lowered by the presence of AI, indicating that AI interaction modes can improve decision clarity. However, the introduction of friction did not lead to increased cognitive load as hypothesised. Additionally, the friction AI modes did not significantly affect task performance compared to frictionless AI modes. This research challenges the notion that reducing user effort maximises efficiency, suggesting that deliberate friction can enhance user engagement and decision-making. Practically, it suggests that AI systems designed with friction elements can improve user decision-making capabilities, making them more effective and user-friendly.

Keywords

AI Design Friction, AI-Human Interactions, Human Centred AI (HCAI), Cognitive State, Usefulness, Confidence, Trust, Sense of Agency, Performance, Generative Artificial Intelligence Assistant (GENAIA).

1.1 Introduction

In an era marked by rapid advancements in artificial intelligence (AI), understanding the dynamics of Human-Artificial Intelligence interaction (HCAI) has become increasingly critical. Traditional HCAI research has predominantly focused on refining the interactions between humans and AI, enhancing efficiency and functionality across various sectors (Xu et al., 2023). Despite these advancements, there remains an important gap regarding the nuanced impact of these interactions on human perceived agency and responsibility. As AI systems, ranging from AI-powered educational tools to algorithms driving financial decisions, become more integrated into daily activities, their role in shaping human decisions and behaviours assumes a greater importance (Holmes & Tuomi, 2022; Hwang & Kim, 2021).

Current HCAI approaches strive for frictionless interactions to maximise efficiency (Mejtoft et al., 2019). However, when combined with AI, this frictionless design approach can lead to negative consequences, such as diminished user agency and responsibility (Pagliari et al., 2022). This raises the need to explore alternative approaches that introduce deliberate 'friction' in AI interactions to enhance user engagement and decision-making (Cox et al., 2016).

To address this gap, this study investigates different modes of AI interactions that aim to increase user's participation, ensuring that technology supports rather than supplants human decision-making. By introducing and evaluating novel AI interaction modes through chatbots built on a generative AI API, referred to as the Generative Artificial Intelligence Assistant (GENAIA), this study aims to understand the user's sense of agency and engagement while maintaining the efficiency of AI-driven tasks.

To examine the impact of these AI interaction modes, this research is guided by the following questions:

- RQ1: To what extent does AI friction affect perceptions towards the assistant and the task performance?
- RQ2: To what extent does AI friction impact the user's objective cognitive state?

These questions aim to understand users' perceptual, emotional, and cognitive responses to varied AI interaction modes and will be addressed through a combination of questionnaires and physiological measurements. This approach allows for a nuanced analysis of the dynamic relation between user perceptions and the physiological sides of human-AI interactions. The study hypotheses that reintroducing a measure of 'friction' in interactions will lead to more enhanced users decision making processes, thus improving user satisfaction and emotional response.

The experiment employed a two-factor mixed design: AI vs. Non-AI as the within-subjects factor and different AI designs as the between-subjects factor. The presence of AI was found to increase participants' confidence and trust while decreasing cognitive load. Within the AI-assisted conditions, different AI interaction designs were tested: a regular chatbot (Non-Friction AI), a chatbot with a summary, and a chatbot with a form (Friction AI). The summary mode significantly enhanced perceived usefulness compared to the other AI modes. Although there were no significant differences in confidence, trust, and task completion time across the AI modes, the summary mode provided a higher sense of agency to users compared to the chatbot with form.

By challenging existing norms of HCAI and proposing a model that balances AI efficiency with human-centric design principles, this research aims to contribute valuable insights into the optimal integration of AI in utilitarian tasks. The theoretical implications extend beyond the current conventions of HCAI by introducing the concept of 'interaction friction,' challenging the prevailing idea that AI system efficiency should be achieved solely through minimal human input. Instead, it posits that deliberate friction can enhance cognitive engagement and decision clarity, fostering a deeper integration of AI in complex decision-making processes.

On a practical level, this research provides actionable insights for the design of AI interfaces. By integrating elements that deliberately increase cognitive friction, designers can create AI systems that support and enhance human decision-making capabilities.

This article is structured as follows: it begins by focusing on the theoretical background of HCAI and AI, followed by the experimental design, data analysis, and results.

1.2 Theoretical Background

1.2.1 Human - AI collaboration

Human-Computer Interaction (HCI) is a key area of study that examines the design and use of computer technology, particularly the interfaces between users and computers. This field includes research on human factors and information systems (Grudin, 2008). HCI researchers both study how humans interact with computers and create new technologies to improve this interaction.

The development of HCI has paralleled technological advancements. Initially, it was a niche field studied mainly by academia and research institutions. The early work in HCI began with Vannevar Bush's concept of the "Memex" in 1945 (Myers, 1998). In the 1960s, Douglas Engelbart and Ted Nelson expanded on Bush's ideas by working on using computers to create and manage complex networks of interlinked text (Baecker, 2008).

Significant progress occurred in the 1960s and 1970s with the development of interactive computer graphics and direct manipulation interfaces (Baecker, 2008). Ivan Sutherland's Sketchpad system showed the potential for effective computer-aided design through innovative graphical user interfaces (GUIs) (Myers, 1998). This era also saw the rise of graphical user interfaces and the WIMP (Windows, Icons, Menus, Pointer) model, which were popularised by personal computers like the Apple Macintosh (Baecker, 2008).

The shift from an academic focus to widespread commercial adoption changed user interaction models, emphasising user satisfaction and interaction efficiency (Grudin, 2008). This shift was a crucial moment in HCI, broadening the focus from usability to include the aesthetic and functional aspects of design, influenced by graphic and industrial designers who developed more user-friendly interfaces (Baecker, 2008).

Computer-Human Interaction (CHI) marks a pivotal evolution from Human-Computer Interaction (HCI), shifting the focus from mere interaction to seamless integration (Xu et al., 2023). With the widespread adoption of personal computers, an increasing demand for improved interfaces and interactions arises. This improvement facilitates technology's integration into daily life, positioning it as an extension of human capabilities rather than just a mere tool. As computers become increasingly embedded across various settings, the interplay between humans and machines could weave so integrally into our lives that it mirrors the ubiquity and unnoticed essentiality of motors in many modern devices (Grudin, 2008).

Artificial intelligence (AI) has been defined as “software (and possibly also hardware) systems designed by humans that, given a complex goal, act in the physical or digital dimension by perceiving their environment through data acquisition, interpreting the collected structured or unstructured data, reasoning on the knowledge, or processing the information, derived from this data and deciding the best action(s) to take to achieve the given goal” (Samoili et al., 2020). AI encompasses a broad spectrum of backend technologies like computer vision and natural language processing (NLP), and user-facing applications such as chatbots and decision support systems (Samoili et al., 2020). AI tools have been implemented in different activity domains in order to increase the efficiency of users. Because modern AI tools are able to automate some routine tasks, the users are able to focus their energy on more challenging problems (Mazingue, 2023). Traditionally to increase efficiency and users satisfaction when interacting with AI, designers have favoured a frictionless approach to AI interactions (Chen & Schmidt, 2024a) However, while this can make processes more efficient, it can also lead to mindless interactions and decisions (Cox et al., 2016). This decrease in users engagement can result in an increase in errors and lack of accountability (Merritt et al., 2019; Yetgin et al., 2015). Balancing users' engagement during an interaction with AI has therefore become an important issue for HCI research (Chen & Schmidt, 2024a).

Human-Centred Artificial Intelligence (HCAI) focuses on developing AI systems that effectively enhance human capabilities and cater to human needs. HCAI integrates human-centred design principles to ensure that AI systems are reliable, understandable, and empowering for users (Ben Shneiderman, 2022). The core principles of HCAI include human factor designs, ethically aligned designs, and technology enhancement (Xu, 2019). These principles ensure that AI systems are effective, operable, and beneficial in real-world settings. By emphasising these principles, HCAI strives to build ethical AI solutions that align with human values. Examples of HCAI in action include AI-driven diagnostic support in healthcare (Jiang et al., 2017), personalised learning systems in education (Holmes & Tuomi, 2022), and AI-enhanced customer services (Chong et al., 2021). These applications demonstrate the potential of HCAI to improve service delivery and user satisfaction across various domains.

1.2.2 Design Friction in AI

Design friction refers to intentional complexities introduced in user interactions to slow down the process, promoting more mindful and reflective decision-making (Mejtoft et al., 2019). This approach contrasts with the prevalent trend of minimising friction to maximise efficiency, which can lead to reduced user agency and over-reliance on AI recommendations (Naiseh et al., 2021). For example, the slow design concept aims to support users in making thoughtful decisions, rather than merely increasing speed and productivity (Mejtoft et al., 2019). Another example of design friction are microboundaries, small obstacles introduced in the interaction flow. They have been shown to help users become more aware of their actions and reduce errors (Cox et al., 2016). Studies have also shown that another form of design friction, nudges, can help users to adjust their trust level in AI, so as to not overly depend on or underlie them (Naiseh et al., 2021). All these designs aim to help users be more mindful of their interaction with AI and improve decision making processes (Cox et al., 2016).

1.2.3 Overview of Chatbots and AI in Human-Computer Interaction

From the first interactions with ELIZA to the emergence of ChatGPT, the landscape of chatbots has greatly evolved over the last decade. Chatbots began as an application of Turing's concerns about whether humans could differentiate interactions between humans and computers, highlighted by the now-famous Turing Test. The first chatbots, such as ELIZA, PARRY, and ALICE, were designed to converse with users and could not perform external tasks (Adamopoulou & Moussiades, 2020). The release of SmartChild and voice assistants marked a shift from chatbots merely conversing to helping users by providing up-to-date information and assisting with schedule management, among other tasks (Adamopoulou & Moussiades, 2020). The emergence of voice assistants also pushed for the development of new designs and interactions with chatbots, including voice, image, and text-based interfaces. Improvements in machine learning and natural language processing (NLP), from supervised to unsupervised and reinforcement learning models, have enhanced the capabilities of chatbots. Chatbots can now serve multiple functions, such as conversational, informative, and task-based (Adamopoulou &

Moussiades, 2020). Task-based chatbots offer many opportunities for companies to streamline some of their processes, a common one is customer service, as chatbots offer 24/7 support to the user, allowing them to ask questions. In the transportation or food industry, chatbots allow for the ordering and tracking of tickets and orders, allowing the employees to focus on more complex tasks.

This study employs three distinct designs of chatbots to explore the effects of different interaction modes on user comprehension, satisfaction, and decision-making efficacy. These variants are designed based on the theoretical frameworks of Media Richness Theory (Kelton et al., 2010), Cognitive Fit Theory (Karran et al., 2022), Split Attention Theory (Pouw et al., 2019) and practical insights derived from recent empirical studies on chatbot interaction.

A. Purpose of AI Text-Based Chatbot

As mentioned, there are different types of chatbots: those that work with customers are called external-facing chatbots, while those that work with employees are referred to as internal-facing chatbots (Gkinko & Elbanna, 2022). In this study, participants were required to interact with an internal-facing chatbot on a company's internal website to perform various utilitarian tasks.

B. Types of AI Text-Based Chatbot

To further identify the impacts of AI interactions on the emotional cognitive and perceptual state of the users different designs of chatbots were used. The section below aims to define and differentiate them.

Frictionless AI Design

The first variant, the Regular Chatbot, serves as the frictionless AI mode of interaction. It follows a standard text-based interaction pattern that has been commonly used in customer service applications (Adamopoulou & Moussiades, 2020). This chatbot operates on a reactive interaction style, responding directly to user inputs. According to Media Richness Theory, this simpler design provides a fundamental baseline for comparing more interactive or visually supported interfaces (Kelton et al., 2010). The richness of a presentation medium is determined by "the ability of information to change understanding within a time interval" (Kelton et al., 2010). This mode allows for an examination of how basic text interactions influence user decision-making and cognitive state in AI-driven chat systems.

Friction AI Design

Two variants of chatbots have been used to integrate friction in the interaction. Other studies have shown that having nudges or microboundaries in AI affected the users trust, satisfaction and decision making processes (Mejtoft et al., 2019; Naiseh et al., 2021).

The first AI friction design features a summary presentation at the top of the chatbot. This mode is based on information presentation research, which suggests that providing a concise summary of crucial information upfront can significantly boost comprehension and decision-making efficiency (Petkovic et al., 2016). From a situation awareness (SA) perspective, this design maintains users' situational awareness by ensuring they can perceive and understand the main elements of the task at hand, which aligns with Endsley's three-tier model of SA. The first level of SA involves the perception of elements in the environment (Endsley, 1995), which the summary presentation facilitates by highlighting key information. The second level of SA, comprehension, is the stage where users are "synthesizing the attributes and dynamics of the identified elements within tier 1" (Jiang et al., 2023). The concise format of the summary allows users to focus on the most important information in the discussion. By enhancing users' situational awareness when interacting with AI, this design improves their confidence and trust in the system. Furthermore, by improving users' situational awareness, the chatbot and summary ensure that users access the information they need, making them more mindful of their decisions (Jiang et al., 2023).

The second Friction AI design is the chatbot with a form. The Chatbot with an integrated form features a dual-pane interface, with the form on the left side of the window and the chatbot on the right. This design draws on the situation awareness theory, as users have access to more elements of information about the task at hand, thereby improving their perception of their environment, the first tier of the situation awareness theory (Endsley, 1995). Nevertheless, this design does not synthesise the information like the summary did, meaning that users have to deal with a lot more information, impacting their comprehension (Tier 2 of situation awareness)

(Jiang et al., 2023). This quantity of information can overwhelm users, or as described in the cognitive load literature, lead to cognitive overload (Fox et al., 2007). In education, cognitive load and multiple panel designs have been related to the split-attention effect, where users struggle to manage their attention across multiple windows (Pouw et al., 2019). This split in attention due to the dual-panel design is therefore expected to lead to different cognitive states and performance outcomes compared to other chatbot designs.

The different literature frameworks allow us to expect some differences between the modes of AI Interaction, however there is yet to have enough literature on the different designs of chatbot to go further in the direction and intensity of the differences.

1.2.4 User Perceptions and Interactions with AI

Artificial Intelligence significantly transforms user experiences and expectations across various domains. AI has reshaped the landscape of user interaction by automating processes and assisting users. This transformation affects users' perception of technology's role in their daily lives and work environments. In turn, these perceptions impact the user's reliance on technological solutions, as highlighted by the Technology Acceptance Model (TAM) (Marangunić & Granić, 2015).

Chatbots, as a widely used form of AI, are conceived to directly interact with the user through different channels, therefore directly impacting their perceptions. These interactions can enhance perceived usefulness, trust, and engagement, particularly when chatbots deliver relevant responses (Kuhail et al., 2022). Completing tasks using a chatbot can also impact the user's sense of agency (Chong et al., 2021). However, the efficacy of these interactions can vary significantly based on the chatbot's design, success, and the complexity of the tasks it manages.

Perceived usefulness has been a recognised measure of perceptions and antecedent of reliance on technology (Davis, 1989). Antecedents of the perceived usefulness of AI include output quality and relevance (Marangunić & Granić, 2015). AI systems that demonstrate significant improvements in task performance or offer considerable convenience are viewed as more beneficial. For utilitarian purposes, it was shown that the more information is provided to the users, the higher the satisfaction of the users and the better the perception of the tool and intention to use it (Cheng & Jiang, 2020). The hypotheses for perceived usefulness include:

- H1a: To perform a utilitarian task, users will find an information system with chatbot assistance more useful than an information system without chatbot assistance.
- H1b: Adding friction to a chatbot will increase its perceived usefulness for performing a utilitarian task. Therefore, chatbots with summary and chatbots with form will be perceived as more useful than the regular chatbot.

Confidence in AI has been defined as "a measure of risk as to how sure users are that they received the correct suggestions by the AI system and if they consider the system to be reliable, i.e., the system consistently operates properly, functional, i.e., the system does what it is supposed to do, and helpful, i.e., the system provides adequate help for the users" (Karran et al., 2022). User confidence in AI depends on the system's performance consistency and its perceived usefulness (Chong et al., 2022; Falconnet et al., 2023). Studies on decision support systems have shown that longer explanations cause more impact on the reliance and intention of the users in accepting recommendations from the decision-making assistant system (Gönül et al., 2006; Wanner et al., 2020). Therefore, it is proposed that:

- H2a: When performing utilitarian tasks, users will exhibit higher confidence in an information system with chatbot assistance compared to an information system without chatbot assistance.
- H2b: Adding friction to a chatbot will increase users' confidence when performing utilitarian tasks. Hence, the modes with friction that present more information than the frictionless AI modes will have a positive effect on users' confidence compared to the frictionless chatbot.

Trust in AI is "the attitude that an agent will help achieve an individual's goals in a situation characterised by uncertainty and vulnerability" (Lee & See, 2004). The system's transparency, reliability, and the accuracy of its outputs have been considered antecedents of trust. Systems with clear, user-friendly interfaces and predictable, consistent behaviours foster greater trust. The availability of additional information about the chatbot has been shown to impact the user's trust in the chatbot (Hudon et al., 2021; Karran et al., 2022). In a study, it was shown that users find systems that require more but shorter steps with the chatbot to be more trustworthy (Casadei et al.). Therefore, we believe that the nodes of interactions that introduce friction will require users to interact more slowly with the assistants to be perceived as more useful and more trustworthy. This study hypotheses:

- H3a: When performing utilitarian tasks, users will exhibit higher trust in an information system with chatbot assistance compared to an information system without chatbot assistance.
- H3b: Adding friction to a chatbot will increase users' trust when performing utilitarian tasks. Therefore, chatbots with summary and chatbots with form will lead to higher levels of declared trust compared to the frictionless chatbot.

The sense of agency has been defined as "the feeling that one is controlling events through one's own behavior" (Wen et al., 2015). The use of AI has been reported to diminish a user's sense of agency in some segments of a task or being removed from the human in favour of automation (Pagliari et al., 2022). Frictionless designs in AI assistants have been accused of increasing users emotional disengagement, decreased agency, and responsibility (Friedland, 2020). On the other hand, it is argued that friction can increase users' sense of self-awareness, prompting them to reflect more on the consequences of their actions (Friedland, 2020). This study proposes:

- H4a: When performing utilitarian tasks, users' sense of agency will be higher in a manual control condition compared to using an information system with chatbot assistance.
- H4b: Adding friction to a chatbot will increase users' perceived sense of agency when performing utilitarian tasks. Therefore, chatbots with summary and chatbots with form will lead to a higher perceived sense of agency compared to the frictionless chatbot.

1.2.5 Cognitive Impact of AI Interaction

The integration of AI into decision-making processes presents some cognitive challenges. The phenomenon of automation bias has been studied for several years in various contexts such as healthcare and transportation (Lyell et al., 2018; Parasuraman & Manzey, 2010). This bias is characterised by an overreliance on AI, such as the cognitive engagement and task outcome are both deteriorating, as well reducing the active involvement of users in decision-making (Pattyn et al., 2008). The risk is also the complete disengagement of the human in the decision processes, research is therefore being made to maintain the human-in-the-loop (Wu et al., 2022). Understanding and measuring cognitive load becomes essential in this context, so as to understand the impact of different modes of AI interaction and ensure balanced collaboration with AI technologies. It was shown in multiple studies how automation can reduce cognitive load (De Bruyne et al., 2023; Young & Stanton, 2002). In "Thinking, Fast and Slow," Dr Kahnman highlights that human decision-making processes employ two systems. The first system is based on intuition and has been described as ‘automatic,’ where no effort is made to come up with the decision. This is the fast system. On the other hand, the second system is considered the slower one, where deliberate decisions and thinking are made (Kahneman, 2011; Shleifer, 2012). In frictionless AI designs, as users become increasingly disengaged with the task and the context, they rely more on their System 1 and do not put as much effort into completing the task (Cox et al., 2016). However, introducing friction, by providing more information for instance, encourages users to be more intentional and active in their decisions, increasing their cognitive activity (Cox et al., 2016).

- H5a: The mode of AI interaction will influence cognitive load during an utilitarian task performance.
- H5b: Adding friction to a chatbot will require more decision making processes. Therefore, chatbots with summary and chatbots with form are expected to cause a higher cognitive load compared to the frictionless chatbot.

1.2.6 Performance Impact of AI Interaction

The popularity of AI assistants and chatbots comes from the expectation of improved performance of the users. Performance in the literature has been conceptualised in different manners (Hemmer et al., 2023). In this study the task performance is operationalized through the metric of time to task completion, which offers a quantifiable measure to assess the efficacy of human-AI collaboration. Studies have shown the positive impacts of AI delegations and interventions on collaborative task performance (Hemmer et al., 2023; Peng et al., 2023). It is hypothesised that task performance time will be reduced using AI interaction modes compared to no AI assistance. Variability in task performance time across different AI modes is also expected, as the presence of friction and added information will lead to the users having to spend more time thinking about their decisions instead of acting in an automatic manner (Chen & Schmidt, 2024b; Kahneman, 2011).

- H6a: Users will perform utilitarian tasks faster when using an information system with chatbot assistance compared to an information system without chatbot assistance.
- H6b: Adding friction to a chatbot will require more steps from the users to perform tasks. Therefore, chatbots with summary and chatbots with form will lead to slower task completion times compared to the frictionless chatbot.

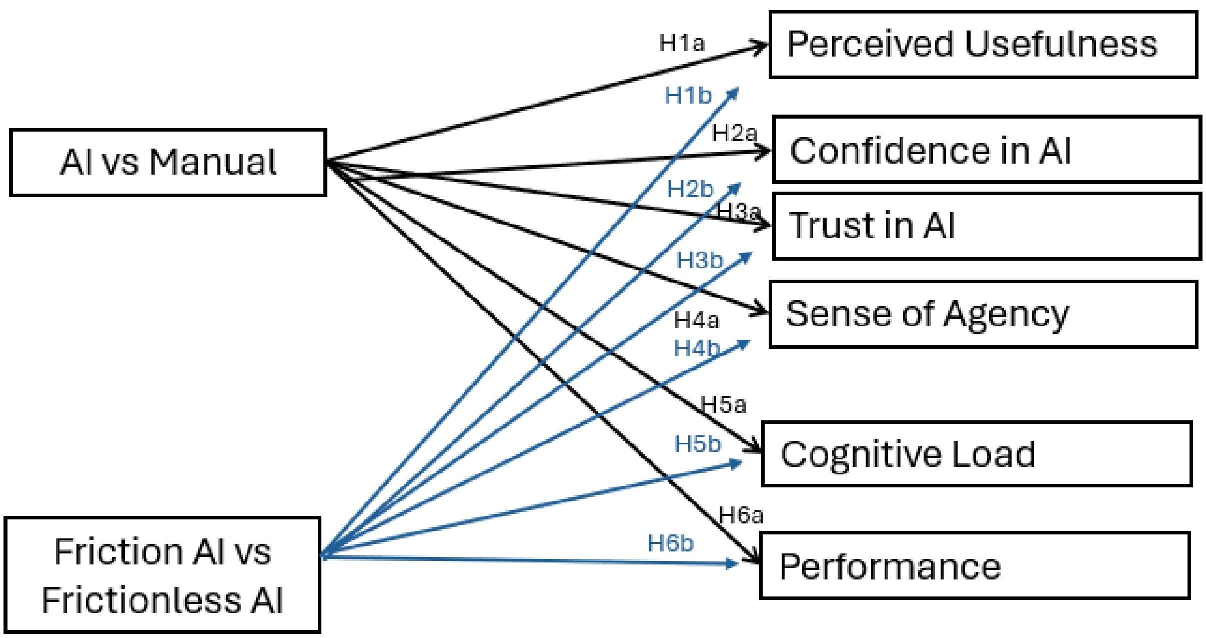

Figure 1: Research Model

1.3 Methodology

Experimental Design

This study aimed to assess the impact of different modes of interaction with artificial intelligence (AI) on participants' perceptions, task performance, and cognitive state. The experimental design was a mixed design: the presence of AI was the within-subjects factor (presence or absence of AI) and the AI interaction mode was the between-subjects factor (Regular AI chatbot, AI chatbot + summary, or AI chatbot + form).

In the control condition, participants performed tasks manually without any AI assistance. These control tasks involved manually finding and completing forms in a paystubs scenario, where participants requested their payslips for a specific period. This served as a baseline for comparing AI-assisted interactions. For the AI-assisted tasks, three distinct modes of interaction were used: a regular text-based chatbot, a chatbot equipped with a summary window on top displaying the requests, and a chatbot that directed participants to forms, allowing interaction with both the chatbot and the form. Participants worked with three different scenarios: in the laptop scenario, users had to order a new computer; in the expense scenario, users had to fill out a form to report a business expense and request a refund; and in the paystubs scenario, users had to request their payslips for a specific period. The order of the tasks involving laptops and expense reports was randomised across participants. Each participant was randomly assigned to one of the three AI interaction modes for the duration of the experiment, ensuring that they experienced both the control condition and one AI-assisted mode.

Sample

The experiment simulated a work environment where 34 individuals recruited from a Canadian Business School were asked to imagine themselves as employees of a company. Eligibility criteria for participants included proficiency in English and basic knowledge of business organisations, along with specific physiological criteria suitable for the measurement tools used.

Ethical considerations were strictly followed, with the study receiving approval from our institution's Research Ethics Board (Ethics Form #2023-5385). Informed consent was obtained from all participants, who were thoroughly informed about the study's nature, their rights as participants, and the confidentiality measures in place.

Stimuli

In this study a Generative Artificial Intelligence Assistant (GENAIA) was used to investigate two factors: AI vs. Non-AI interactions and Non-Friction AI vs Friction AI. Two designs of Friction AI were used, the first one displayed a summary of the interactions above the chatbot window and the second one showed the form the participants had to fill alongside the chatbot. The assistant was used in a utilitarian context on a prototype of a company's intranet. In the control condition, where users had to perform tasks manually, the website was identical but the AI was deactivated.

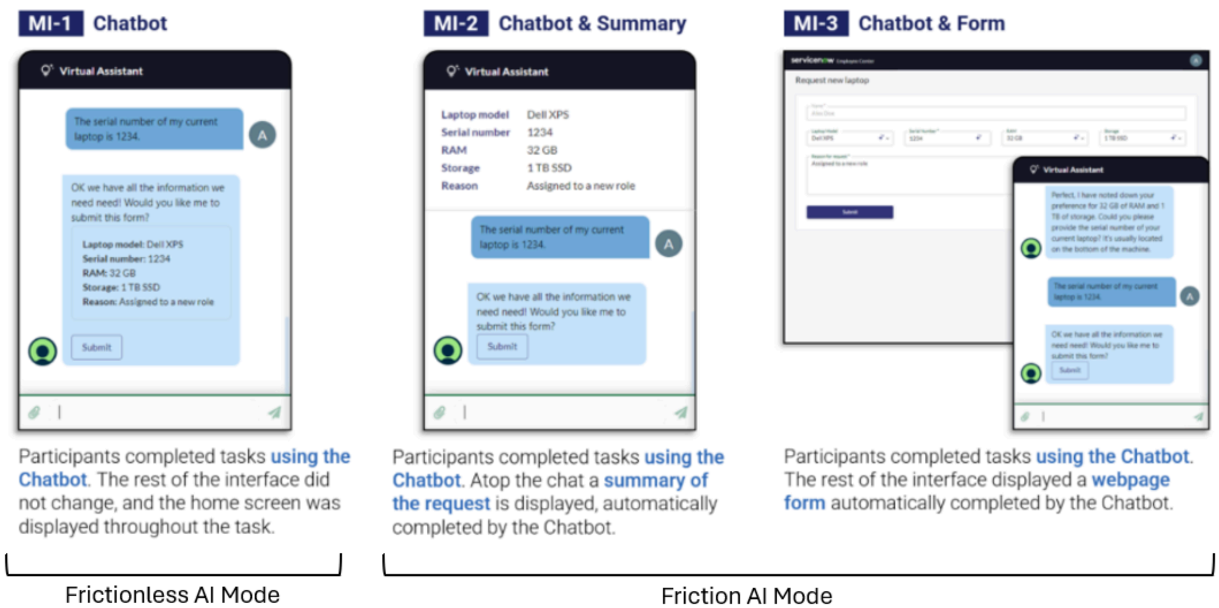

Figure 2: Different modes of chatbots

Procedure

Participants began the experiment by signing a consent form, after which physiological measurement tools were installed. They then focused on a cross displayed on the screen for 90 seconds to obtain a baseline measurement. This procedure was designed to stabilise physiological measures, providing a default physiological test prior to interacting with the stimuli. Each participant then engaged in a series of eight tasks, divided into control tasks without AI assistance and AI-assisted scenarios. Participants were randomly assigned to one of three AI interaction design conditions: no AI, basic AI, or advanced AI. The assignment to these conditions was randomised to ensure equal distribution of participants across the three groups, and each participant experienced only one AI condition throughout the experiment. We assumed that no cross-over effects occurred between conditions as participants were exposed to only one AI condition, preventing any contamination between tasks.

The scenarios within each condition were also randomised, with participants exposed to different task scenarios in a random order. This included randomization of Scenarios 2 and 3, which were designed to vary between participants, while Scenario 1 remained consistent and was always presented first to establish a baseline. In all cases, participants always began with a control (no AI) task, after which the randomised AI-assisted tasks were introduced.

A pre-study questionnaire administered via Qualtrics gathered demographic information and assessed participants' familiarity and self-efficacy with AI. This questionnaire aimed to establish a contextual baseline for subsequent analyses. To measure cognitive load, an eye tracker placed below the participant's monitor was used to record pupil dilation (Beatty, 1982). After each task, participants completed additional questionnaires measuring confidence in AI, trust in AI, sense of agency, and perceived usefulness. Due to the repetitive nature of the questionnaires, attention checks were included between some questions to ensure participants were reading the questions properly. After completing all tasks and questionnaires, participants participated in an interview to provide further insights into their experiences during the tasks.

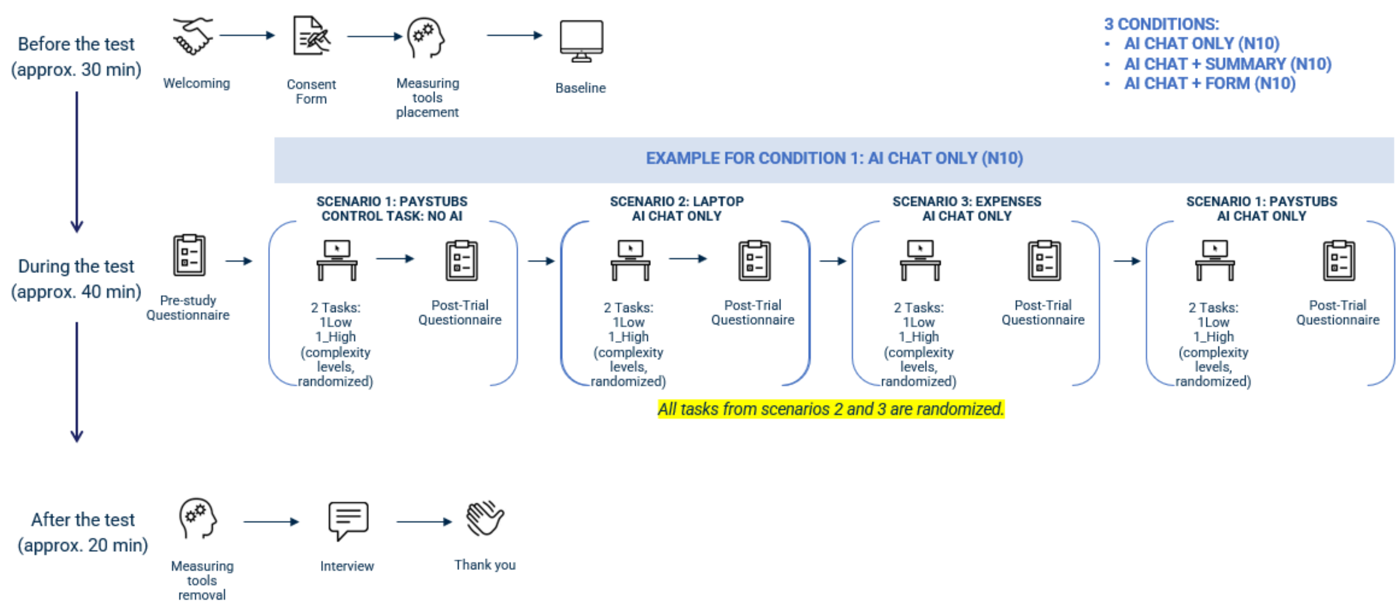

Figure 3: Experimental Design

Instruments and Measures

The data collection process involved two main tools. The measurement scales were taken from the literature. The participants were asked to evaluate on a likert scale of 1 to 7 their degree of confidence in the output, their degree of trust in the system (Falconnet et al., 2023) and their perceived usefulness (Davis, 1989). The participants were also asked to rate on a scale of 1 to 100 their perceived sense of control on the output during the task (Metcalfe & Greene, 2007). This physiological measure provided an inference of cognitive load, with adjustments made based on the initial 90-second baseline recording. The setup allowed for an objective measurement of the cognitive state changes in response to different interaction scenarios.

One of the aims of this research project is to understand the impact of different modes of interaction with AI on the user's cognitive load. Based on the literature (Beatty, 1982), one of the tools used to measure this cognitive load was an eye tracker. This method is commonly used as a non-invasive procedure to measure cognitive load by comparing participant's pupil diameter during a task with a baseline (Krejtz et al., 2018). In this study, Tobii Pro Lab eye tracking software and hardware were used. The system records images of the eyes and uses algorithms to generate 3D models of the participant's eyes (Tobii, 2023). Data was recorded in milliseconds, with a sampling interval of 16.87 milliseconds and a frequency of 59.3 Hz (rounded to 60 Hz) (Tobii, 2024).

Central to this study is the hypothesis that tasks performed with AI assistance will exhibit a different level of cognitive workload compared to those executed without AI assistance (Brachten et al., 2020). This variance is anticipated to be reflected in the degree of change in pupil size from the established baseline, indicating either an increase or decrease in cognitive workload.

A fundamental aim of this study is to examine the efficiency of tasks performed under varied modes of AI interaction as it relates to user performance in a controlled setting. The dataset derived from these trials includes metrics such as task identifiers, modes of AI interactions, participant IDs, and the time taken for task completion across 246 observations. The time metric was extracted from the Tobii software using the start of task and end of task markers, by default they were recorded in milliseconds.

Construct	Measurement	Measurement tool	Reference
Perceived Usefulness	Scale (1-7)	Using this system in my job would enable me to accomplish tasks more quickly. Using this system would improve my job performance. Using this system in my job would increase my productivity. Using this system would enhance my effectiveness on the job. Using this system would make it easier to do my job.	(Davis, 1989)
Sense of Agency	Scale (1-100)	How much in control were you during this task?	(Metcalfe & Greene, 2007)
Confidence in AI	Scale (1-7)	I am convinced of the output given to me by the system.	(Falconnet et al., 2023)
Trust in AI	Scale (1-7)	The system can be trusted.	(Falconnet et al., 2023)
Cognitive Load	Pupil dilatation (mm)	Eye Tracker Tobii Pro Lab (Tobii AB, Stockholm, Sweden)	(Beatty, 1982)
Performance	Time to task		(Chen & Schmidt, 2024b)

Table 2 : Measurement Tools

1.4 Data Analysis

Measurement Scale Analysis

In this study, a significant focus was placed on understanding participants' perceptions of the AI assistant, specifically exploring their attitudes and reactions during various tasks. Participants were asked to evaluate their perception of the AI assistant, particularly in terms of trust, confidence, and perceived usefulness. In addition, measurement scales were used to analyse the participants' perceptions of their sense of agency. This following section will detail the methodologies employed in processing and analysing this data.

Preprocessing data obtained from Qualtrics questionnaires is an important step in ensuring the reliability and validity of subsequent analyses. This process, conducted using Python (3.8.13), involves the libraries Pandas (1.4.4), NumPy (1.21.5), and Matplotlib (3.5.3). The pre-processing streamlines data cleaning across all participants by selecting relevant columns and removing initial non-participant rows to reset the index for consistency. It further includes reshaping the data to organise it by unique participants and scenarios according to the independent variables, ensuring a comprehensive dataset reflecting various dimensions of the survey. Additional cleaning processes merge and standardise data types across different data frames and incorporate a validity dataframe based on experimental observations, such as excluding participants affected by technical issues, to maintain the integrity of the analysis. Observations for which participants failed the attention check were also removed.

An initial descriptive analysis was performed to verify the assumptions regarding the distribution of the data. Observations revealed that the data for all variables did not follow a normal distribution, prompting the decision to pursue a nonparametric approach. As such, Kruskal-Wallis tests were used to detect statistically significant differences across manipulation intensities, as this test is appropriate for non-normally distributed data and comparisons between three or more independent groups. Kruskal-Wallis was applied to each perceived dependent variable, such as perceived usefulness, sense of agency, trust, and confidence. Post-hoc analysis using Dunn's tests with Bonferroni correction was implemented to further explore significant differences between interaction modes, adjusting for the increased risk of Type I errors in multiple comparisons. In addition to ensure the robustness of the analysis, dummy variables were created for the independent variable Modes of Interaction and the Variance Inflation Factor (VIF) values were calculated for these dummy variables. All VIF values were found to be below the threshold of 5, indicating that multicollinearity is not a significant concern for these variables.

Additionally, Spearman's rank correlation analysis was performed to explore the relationships between the dependent variables. This analysis revealed significant correlations, providing further insights into how these variables interact with each other.

Eye Tracker

The methodology employed for preprocessing leverages Python (3.8.13) with packages such as Pandas (1.4.4), NumPy (1.21.5), and Matplotlib (3.5.3), ensuring efficient data management and preprocessing. The initial step in the data processing was a descriptive examination, providing a preliminary statistical overview and presenting data distribution through scatter plots and Q-Q plots. Once again multicollinearity was tested using the dummy variables previously created and the results were all below the threshold of 5. The spearman's rank correlation showed that the eye tracker data was not significantly correlated with other dependent variables. This examination inferred the usability of the eye tracker for 34 participants and confirmed that the pupil data follows a normal distribution, with scatter plots indicating no multicollinearity issues.

The task segmentation strategy focused on determining the minimum time required for the completion of each task. By normalising the task segments using this minimum time frame and recognizing keyboard interactions within these segments, the approach allowed for a consistent comparison across all participants.

The inter-baseline pupil size was calculated by measuring the average pupil size during a predetermined baseline interval of 90 seconds, simplifying the computation by averaging the records from both eyes (Menekse Dalveren & Cagiltay, 2019). The data was then cleansed to ensure high validity scores, categorising tasks into AI-supported and non-AI groups. Participants with invalid baselines were removed, as were observation segments not belonging to the baseline or a task. Keyboard events were also categorised, with some events labelled as 'pulse' based on specific criteria related to the device used.

A segment was created for each task by identifying the first keyboard interaction after the task's start. This involved parsing participant data to locate these events with their corresponding timestamps and delineating the segments accordingly. Each segment ranged from 1000 milliseconds before to 1500 milliseconds after the first keyboard interaction, capturing a comprehensive snapshot of the participant's physiological reaction during that period. This segment allowed us to measure and compare the average pupil diameter across different conditions. Once extracted, the average pupil size for both eyes was computed. To account for individual participant characteristics, the difference between the baseline state and the segment was calculated, providing the average change in pupil size due to the stimuli interaction.

The final stages of the methodology involve a comprehensive statistical analysis using R (4.2.1), where the segmented data is imported for processing.

The primary tool for this analysis was the Linear Mixed Models (LMM), which were applied to accommodate the multi-layered data structure, incorporating both fixed effects with the modes of Interaction and random effects attributable to individual participant differences (Magezi, 2015).

To further refine the analysis, Analysis of Variance (ANOVA) tests were employed. These tests were crucial in determining the statistical significance of the differences in cognitive load observed across the AI interaction modes (Starkweather, 2010). Conducted for each segment of the task, ANOVA was used to detect significant differences between interaction modes because the data for these variables met the assumption of normality, as confirmed by initial descriptive statistics and Q-Q plots.

In addition to LMM and ANOVA, a post-hoc analysis, the Tukey HSD (Honestly Significant Difference) test, was performed in the statistical analysis following the ANOVA. This test was used to identify specific pairwise differences between the modes of AI interaction, providing detailed insights into which specific modes significantly differed from each other in terms of cognitive load (Abdi & Williams, 2010). The use of the Tukey HSD test thus offered a more granular understanding of the statistical significance of the observed differences.

Performance Analysis

The analysis began with an initial descriptive examination of the dataset to verify its structure. The data for task time was notably right-skewed, so a log transformation was applied to normalise the distribution. The VIF scores for task time were all below the threshold of 5, and the Spearman's rank correlation didn't reveal significant correlations with other dependent variables.

Linear Mixed Models (LMM) were employed to account for both fixed and random effects within the dataset. A random effect for the participants was included to generalise the impact of AI interaction modes on the log-transformed task times, minimising the noise from individual participant characteristics. After fitting these models, estimated marginal means (EMMs) were computed to estimate the adjusted average task completion times for each interaction mode, taking into account other factors in the model, such as participant effects (Searle et al., 1980).

Pairwise comparisons of these EMMs were then conducted to identify significant differences between each mode of interaction. The results of these comparisons were adjusted for multiple testing using the Bonferroni correction, addressing the increased risk of Type I errors that come with multiple comparisons.

1.5 Results

Descriptive statistics

The table 3 below provides a summary of the descriptive statistics for the dependent variables in this study, including Perceived Usefulness, Confidence in AI, Trust in AI, Sense of Agency, Pupil Size (as a measure of cognitive load), and Completion Time (as a measure of Performance). The table presents the means (M) and standard deviations (SD) for each variable across different design conditions: Chatbot (Frictionless Design), Chatbot with Summary, Chatbot with Form, and Without AI (Control), alongside the statistics for the full sample.

Variable	Chatbot		Chatbot + Summary		Chatbot + Form		Without AI (Control)		Full Sample	
Statistic	M	SD	M	SD	M	SD	M	SD	M	SD
Perceived Usefulness	5.47	1.19	5.88	1.07	5.34	1.28	5.05	1.06	5.44	1.19
Confidence in AI	5.94	0.97	6.17	0.76	5.82	1.35	5.4	1.06	5.83	1.10
Trust in AI	6.10	0.714	6.02	0.911	5.62	1.47	5.31	1.15	5.751	1.16
Sense of Agency	76.70	25	77.7	29.80	70.6	20.60	77.70	19.30	75.48	24.05
Pupil Size	0.45	0.29	0.44	0.41	0.57	0.44	0.80	0.39	0.57	0.41
Completion Time (ms)	159064	85132	130996	66057	148227	84045	123995	83883	139774	80618
Legend	M = Mean SD = Standard Deviation									

Table 3: Descriptive Statistics

Correlation Analysis

Table 4 presents the Spearman correlation coefficients between the dependent variables in this study, including Perceived Usefulness (PUS), Confidence in AI (CAI), Trust in AI (TAI), Sense of Agency (AGY), Pupil Size (as a measure of cognitive load), and Completion Time (as a measure of performance).

Given that participants completed multiple tasks, repeated measurements were averaged across tasks for each participant to ensure that the correlation analysis reflects a single value per participant, controlling for intra-individual variation. This approach ensures that the correlation coefficients capture overall trends across participants rather than task-specific fluctuations.

A strong positive correlation is observed between Trust in AI and Confidence in AI ($\rho = 0.766$), indicating that higher confidence in AI is associated with increased trust. Perceived Usefulness is also correlated with both Trust in AI ($\rho = 0.589$) and Confidence in AI ($\rho = 0.543$), suggesting that participants who perceive the AI as more useful also tend to have greater confidence and trust in it.

On the other hand, some variables, such as Pupil Size (Cognitive Load), Completion Time (Performance), and Sense of Agency, show weak or non-significant correlations with others, indicating minimal or no direct relationship within this context.

	Perceived Usefulness	Confidence in AI	Trust in AI	Sense of Agency	Pupil Size	Completion Time
Perceived Usefulness	1	0.54	0.59	0.17	0.04	0.06
Confidence in AI	0.543	1	0.766	0.049	-0.051	0.160
Trust in AI	0.59	0.77	1	0.26	-0.03	0.17
Sense of Agency	0.17	0.05	0.26	1	-0.01	-0.06
Pupil Size	0.04	-0.05	-0.03	-0.01	1	0.01
Completion Time	0.06	0.16	0.17	-0.06	0.01	1

Table 4: Spearman's Correlation Matrix

Hypothesis testing

Perceived Usefulness (H1)

In examining the perceived usefulness of AI interaction modes compared to a control condition (H1a), the results indicated a significant difference in perceived usefulness scores across the modes of AI, $\chi^2(3) = 17.697$, $p = .001$, supporting H1a that AI interaction modes are perceived as more useful than the control condition where the users had to fill the form without any assistance.

Further analysis to explore differences in perceived usefulness among various AI interaction modes introducing frictions or not (H1b) revealed significant differences. The comparison between the chatbot with summary and the chatbot with form showed that the presence of a summary was perceived as more useful than the design with form (p-adjusted = .0264). However, no additional difference was found regarding the frictionless regular chatbot mode. The hypothesis H1b cannot be supported since there is no significant result for the comparison between the AI modes with friction and the frictionless mode of interaction. However, this does not indicate that there is no effect, but rather that the evidence was insufficient to reject the null hypothesis.

Confidence in AI (H2)

Regarding the hypothesis that users will exhibit higher confidence in AI interaction modes compared to control (H2a), the Kruskal-Wallis rank sum test revealed a significant difference in confidence across the groups, $\chi^2(3) = 17.496$, $p < 0.001$. The results indicated significantly higher confidence in the AI interaction modes compared to the control condition. Notably, the confidence levels for the frictionless regular chatbot mode compared to control (p-adjusted = 0.019), the chatbot with summary mode compared to control (p-adjusted < 0.001), and the chatbot with form mode compared to control (p-adjusted = 0.005) were all significantly higher, supporting the H2a hypothesis that AI interaction enhances user confidence in task performance.

For the hypothesis regarding the addition of friction in an interaction and its increase on users confidence (H2b), the results showed no significant differences in the adjusted p-values for comparisons between the different AI interaction modes themselves, such as between the frictionless regular chatbot mode and the friction AI modes (p-adjusted = 1.000 the comparison between regular chatbot and chatbot with summary and regular chatbot and chatbot with form) and between the chatbot with summary and chatbot with form. This indicates that while AI modes are associated with higher confidence than the control, friction in AI interaction doesn't induce an increase in the user's confidence compared to the frictionless AI modes of interaction. Therefore, H2b is not supported. The lack of significance does not imply that friction has no effect, but that the evidence was insufficient to reject the null hypothesis.

Trust in AI (H3)

Addressing Hypothesis 3a that users will exhibit higher trust in AI interaction modes compared to the control, the Kruskal-Wallis rank sum test confirmed significant differences in trust levels across different interaction modes and control, $\chi^2(3) = 16.259$, $p = 0.001$. Specifically, the test results showed significantly higher trust for the AI modes compared to the control, as indicated by the comparisons: chatbot-only mode versus control, p-adjusted = 0.001; chatbot with summary mode (MI1) versus control, p-adjusted = 0.002; and chatbot with form mode versus control, p-adjusted = 0.047. The hypothesis H3a that the presence of AI would increase the user's trust in the system is therefore supported.

Regarding Hypothesis H3b, which posits that the presence of friction in the interaction with AI would increase the user's trust in the system, the analysis showed no significant differences in trust levels among the AI interaction modes themselves. The adjusted p-values for these internal AI mode comparisons (regular-chatbot - chatbot-summary, regular-chatbot - chatbot-form) all exceeded the threshold for significance ($\alpha = 0.05$), suggesting that while trust levels were generally higher for AI modes compared to the control, the addition of AI friction interaction did not significantly affect the level of trust. H3b is not supported by the results, as there were no significant variations in trust levels between the frictionless and friction AI modes, indicating that the presence of AI, in general, increases trust rather than specific features or types of AI interaction.

Sense of Agency (H4)

Addressing Hypothesis 4a, which suggests that users' sense of agency will be higher when performing tasks manually in a control condition than with AI, the Kruskal-Wallis rank sum test results indicate that some differences in the user's sense of agency are observed due to the different modes of interaction, $\chi^2(3) = 9.524$, $p = 0.023$. However, the comparison of sense of agency between the AI modes and the control condition did not show significant differences as indicated by the adjusted p-values (p-adjusted = 1.000 for regular chatbot - no ai, p-adjusted = 0.854 for chatbot-summary - no ai, and p-adjusted = 0.188 for chatbot-form -no ai). These results suggest that the sense of agency reported by users was not significantly higher in the control condition compared to the AI interaction modes. While H4a is not supported, this only means that the data did not provide sufficient evidence to reject the null hypothesis; it does not imply that there is no effect.

Concerning Hypothesis 4b, which expects variability in the sense of agency across between the frictionless AI mode and the ones including friction, the test results revealed a significant difference between some of the AI modes. Specifically, the comparison between chatbot with summary mode and chatbot with form mode yielded a significant result, p-adjusted = 0.008 with a Z score of 3.021, meaning that the median perceived sense of agency was higher for the mode with summary than the mode with form. This suggests that there is indeed some variability in the sense of agency within the friction AI modes. However, other comparisons with the frictionless AI modes did not show significant differences (p-adjusted > 0.2), indicating that not all AI interaction modes significantly affect the sense of agency. H4b is therefore not supported.

Cognitive Load (H5)

Addressing Hypotheses 5a and 5b, which suggest that cognitive load will be significantly influenced by the mode of AI interaction, and that the variations of cognitive load are expected among the different modes of interaction. The Kruskal-Wallis rank sum test results indicated a significant difference in cognitive load across the various modes of AI interaction, $\chi^2(3) = 32.79$, $p < 0.001$.

Subsequent pairwise comparisons using Bonferroni correction revealed that cognitive load in the control mode was significantly different from modes with a chatbot-only, chatbot and summary, and chatbot with form, all $p < 0.05$ with positive z scores indicating a higher median level for the control mode. Therefore supporting H5a, indicating that cognitive load is significantly influenced by the presence of AI. However, no significant differences were found between the frictionless regular chatbot and friction modes chatbot-summary, modes chatbot-only and chatbot-form, or modes chatbot-summary and chatbot-form. Hence, the hypothesis H5b was not supported, as no significant variations in cognitive load were identified within the modes of AI interaction. The lack of significant results means that the evidence was insufficient to reject the null hypothesis, not that there is no effect.

Performance (H6)

The hypothesis H6a posits that the presence of AI will significantly reduce the time required to complete tasks. Moreover the hypothesis H6b adds that some variation in time task completion is expected among the different modes of AI interaction.

A series of linear mixed-effects models were fitted to evaluate the impact of different AI interaction modes on task performance time. The results showed that the time to task was multiplied by $\exp(0.2924) = 1.34$ milliseconds when the participants completed the task with the frictionless regular chatbot mode compared to the manually without AI assistance ($SE= 0.106$, $p= 0.032$). However, no significant differences in task performance time were found within the friction AI modes (chatbot and summary and chatbot with form) and the control mode. Therefore the hypothesis H6a is not supported. Hypothesis H6b was also not supported, no time to task differences were found to be significant between the frictionless and friction AI modes. The Spearman's rank correlation didn't reveal significant correlations with other dependent variables, suggesting that the performance measure is relatively independent of the perceptions and cognitive load variables.

Hypotheses results summary

	Hypothesis	Supported or Not
H1a	To perform a utilitarian task, users will find an information system with chatbot assistance more useful than an information system without chatbot assistance.	Supported
H1b	Adding friction to a chatbot will increase its perceived usefulness for performing a utilitarian task. Therefore, chatbots with summary and chatbots with form will be perceived as more useful than the regular chatbot.	Not Supported
H2a	When performing utilitarian tasks, users will exhibit higher confidence in an information system with chatbot assistance compared to an information system without chatbot assistance.	Supported
H2b	Adding friction to a chatbot will increase users' confidence when performing utilitarian tasks. Hence, the modes with friction that present more information than the frictionless AI modes will have a positive effect on users' confidence compared to the frictionless chatbot.	Not Supported
H3a	When performing tasks, users will exhibit higher trust in an information system with chatbot assistance compared to an information system without chatbot assistance.	Supported
H3b	Adding friction to a chatbot will increase users' trust when performing utilitarian tasks. Therefore, chatbots with summary and chatbots with form will lead to higher levels of declared trust compared to the friction-less chatbot.	Not Supported
H4a	When performing utilitarian tasks, users’ sense of agency will be higher in a manual control condition compared to using an information system with chatbot assistance.	Not Supported
H4b	Adding friction to a chatbot will increase users' perceived sense of agency when performing utilitarian tasks. Therefore, chatbots with summary and chatbots with form will lead to a higher perceived sense of agency compared to the frictionless chatbot.	Not Supported
H5a	The mode of AI interaction will significantly influence cognitive load during task performance.	Supported
H5b	Adding friction to a chatbot will require more decision making processes. Therefore, chatbots with summary and chatbots with form are expected to cause a higher cognitive	Not Supported

	load compared to the frictionless chatbot.	
H6a	Users will perform tasks faster when using an information system with chatbot assistance compared to an information system without chatbot assistance.	Not Supported
H6b	Adding friction to a chatbot will require more steps from the users to perform tasks. Therefore, chatbots with summary and chatbots with form will lead to slower task completion times compared to the frictionless chatbot.	Not Supported

Table 4: Hypotheses Results Summary

1.6 Discussion

Summarising Key Findings

This study explored the impact of adding interaction friction in AI-human interactions, using chatbots built on a Generative Artificial Intelligence Assistant (GENAIA). The key findings indicate that AI interaction modes were generally perceived as more useful compared to non-AI conditions. Users reported higher levels of confidence and trust in AI-assisted interactions. Interestingly, the introduction of interaction friction did not significantly enhance perceived usefulness, confidence, trust, or task performance compared to frictionless AI modes. Cognitive load was influenced by the presence of AI, with AI interactions reducing cognitive load compared to the control condition.

Interpreting Results

The findings align with the current movement that argues that while traditional HCI approaches strive for frictionless interactions to maximise efficiency, they may inadvertently diminish user agency and responsibility (Cox et al., 2016). This study expands on this notion by demonstrating that while the presence of AI seems to increase users' trust, confidence, and sense of agency, the addition of friction did not negatively affect users' perceptions. Moreover, the results challenge the 'Don't Make Me Think Movement' (Krug, 2005), which implies that frictionless designs improve users' experience and performance. Instead, the data suggest that introducing friction can maintain, if not improve, decision clarity and user engagement.

Theoretical contributions

This study adds to the understanding of controlled interaction friction in AI systems by studying its nuanced role in enhancing user experience while balancing cognitive load. Previous literature has established the benefits of frictionless AI interactions in improving efficiency and user satisfaction (Chen & Schmidt, 2024b). However, it has also highlighted the risks of such designs leading to reduced user engagement and potential errors (Cox et al., 2016; Merritt et al., 2019). This study findings build on this work by showing that controlled interaction friction can offer a middle ground where cognitive load is reduced, yet users remain actively engaged in the decision-making process.

Specifically, this study contributes to the ongoing research on design friction by confirming the theoretical arguments posited by Mejtoft et al. (Mejtoft et al., 2019) and Naiseh et al. (Naiseh et al., 2021), who suggested that intentional design complexities, such as microboundaries and nudges, could promote more mindful interactions. This study adds to the arguments by demonstrating that users experience lower cognitive load with AI assistance compared to manual task performance, even when friction elements are introduced. This finding differs from the common assumption that friction inherently increases cognitive effort and supports the notion that well-designed friction can enhance situational awareness and decision-making without overwhelming the user (Endsley, 1995; Jiang et al., 2023).

Moreover, the study shows that users reported higher levels of control, confidence, and trust in AI systems even when friction elements, such as summaries and forms, were integrated into the system. This challenges the frictionless approach advocated in traditional HCI literature, where the primary focus has been on minimising cognitive load to enhance user experience (Grudin, 2008). The findings align with and expand the perspectives of (Mejtoft et al., 2019), who emphasised that slowing down interactions through design friction can lead to more deliberate and thoughtful user decisions.

By linking these outcomes to established theories like Media Richness Theory (Kelton et al., 2010)and Situation Awareness Theory (Endsley, 1995), this study supports existing models and shows that controlled friction can simultaneously reduce cognitive load and improve user control.

In summary, this study contributes to the literature by illustrating how controlled interaction friction can serve as a strategic design element to balance cognitive load and user engagement. These findings advance the theoretical understanding of friction in AI-human interaction and offer practical implications for designing AI systems that are both efficient and user-centric.

Practical Contributions

From a practical standpoint, this study offers actionable insights for AI interface design. By incorporating friction elements that intentionally require users to maintain an active role in the interaction, designers can create AI systems that better support and enhance human decision-making capabilities. The strong correlations between perceived usefulness, confidence, and trust suggest that AI interface designs should focus on enhancing these perceptions due to their interrelated nature.

Limitations and research opportunities

Despite its contributions, this study has several limitations. The sample size was relatively small, which may limit the generalisation of the findings. Additionally, the experimental design focused on two specific AI friction designs, which does not capture the full range of possible friction designs,such as pop-ups that encourage users to pause and reflect, or designs that integrate the rationale behind AI recommendations (Chen & Schmidt, 2024b; Naiseh et al., 2021). In addition, this research focused solely on utilitarian tasks. While this study focused on utilitarian tasks, future research should explore larger and more diverse scenarios and friction designs. Researching other forms of AI interaction friction and their effects on different user populations and task contexts would also be valuable.

References List Article 1

Abdi, H., & Williams, L. J. (2010). Tukey's honestly significant difference (HSD) test. Encyclopedia of research design, 3(1), 1-5.

Adamopoulou, E., & Moussiades, L. (2020). Chatbots: History, technology, and applications. Machine Learning with Applications, 2, 100006. https://doi.org/https://doi.org/10.1016/j.mlwa.2020.100006

Baecker, R. M. (2008). TIMELINES Themes in the early history of HCI---some unanswered questions. Interactions, 15(2), 22-27.

Beatty, J. (1982). Task-evoked pupillary responses, processing load, and the structure of processing resources. Psychological Bulletin, 91(2), 276-292. https://doi.org/10.1037/0033-2909.91.2.276

Brachten, F., Brünker, F., Frick, N. R. J., Ross, B., & Stieglitz, S. (2020). On the ability of virtual agents to decrease cognitive load: an experimental study. Information Systems and e-Business Management, 18(2), 187-207. https://doi.org/10.1007/s10257-020-00471-7

Casadei, A., Schlogl, S., & Bergmann, M. (2022). Chatbots for Robotic Process Automation: Investigating Perceived Trust and User Satisfaction.

Chen, Z., & Schmidt, R. (2024a). Exploring a Behavioral Model of "Positive Friction" in Human-AI Interaction. International Conference on Human-Computer Interaction,

Chen, Z., & Schmidt, R. (2024b). Exploring a Behavioral Model of "Positive Friction" in Human-AI Interaction. In (pp. 3-22). Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-61353-1_1

Cheng, Y., & Jiang, H. (2020). How Do AI-driven Chatbots Impact User Experience? Examining Gratifications, Perceived Privacy Risk, Satisfaction, Loyalty, and Continued Use. Journal of Broadcasting & Electronic Media, 64(4), 592-614. https://doi.org/10.1080/08838151.2020.1834296

Chong, L., Zhang, G., Goucher-Lambert, K., Kotovsky, K., & Cagan, J. (2022). Human confidence in artificial intelligence and in themselves: The evolution and impact of confidence on adoption of AI advice. Computers in Human Behavior, 127, 107018. https://doi.org/https://doi.org/10.1016/j.chb.2021.107018

Chong, T., Yu, T., Keeling, D. I., & de Ruyter, K. (2021). AI-chatbots on the services frontline addressing the challenges and opportunities of agency. Journal of Retailing and Consumer Services, 63. https://doi.org/10.1016/j.jretconser.2021.102735

Cox, A. L., Gould, S. J., Cecchinato, M. E., Iacovides, I., & Renfree, I. (2016). Design frictions for mindful interactions: The case for microboundaries. Proceedings of the 2016 CHI conference extended abstracts on human factors in computing systems,

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

De Bruyne, J., Joundi, J., Morton, J., Zheleva, A., Van Kets, N., Van Wallendael, G.,…Bombeke, K. (2023). I spy with my AI: The effects of AI-based visual cueing on human operators' performance and cognitive load in CCTV control rooms. International Journal of

Industrial Ergonomics, 95, 103444. https://doi.org/https://doi.org/10.1016/j.ergon.2023.103444

Endsley, M. R. (1995). Toward a Theory of Situation Awareness in Dynamic Systems. Human Factors: The Journal of the Human Factors and Ergonomics Society, 37(1). https://doi.org/info:doi/10.1518/001872095779049543

Falconnet, A., Coursaris, C. K., Beringer, J., Van Osch, W., Sénécal, S., & Léger, P.-M. (2023). Improving User Experience with Recommender Systems by Informing the Design of Recommendation Messages. Applied Sciences, 13(4), 2706. https://doi.org/10.3390/app13042706

Fox, J. R., Park, B., & Lang, A. (2007). When Available Resources Become Negative Resources:The Effects of Cognitive Overload on Memory Sensitivity and Criterion Bias. Communication Research, 34(3), 277-296. https://doi.org/10.1177/0093650207300429

Friedland, J. (2020). AI can help us live more deliberately.

Gkinko, L., & Elbanna, A. (2022). Hope, tolerance and empathy: employees' emotions when using an AI-enabled chatbot in a digitalised workplace. Information Technology & People, 35(6), 1714-1743. https://doi.org/10.1108/itp-04-2021-0328

Grudin, J. (2008). A moving target: The evolution of HCI. The human-computer interaction handbook: Fundamentals, evolving technologies, and emerging applications, 1-24.

Gönül, M. S., Önkal, D., & Lawrence, M. (2006). The effects of structural characteristics of explanations on use of a DSS. Decision support systems, 42(3), 1481-1493.

Hemmer, P., Westphal, M., Schemmer, M., Vetter, S., Vössing, M., & Satzger, G. (2023, 2023). Human-AI Collaboration: The Effect of AI Delegation on Human Task Performance and Task Satisfaction.

Holmes, W., & Tuomi, I. (2022). State of the art and practice in AI in education. European Journal of Education, 57(4), 542-570.

Hudon, A., Demazure, T., Karran, A., Léger, P.-M., & Sénécal, S. (2021). Explainable Artificial Intelligence (XAI): How the Visualization of AI Predictions Affects User Cognitive Load and Confidence. In (pp. 237-246). Springer International Publishing. https://doi.org/10.1007/978-3-030-88900-5_27

Hwang, S., & Kim, J. (2021). Toward a Chatbot for Financial Sustainability. Sustainability, 13(6), 3173.

Jiang, F., Jiang, Y., Zhi, H., Dong, Y., Li, H., Ma, S.,…Wang, Y. (2017). Artificial intelligence in healthcare: past, present and future. Stroke and vascular neurology, 2(4).

Jiang, J., Karran, A. J., Coursaris, C. K., Léger, P.-M., & Beringer, J. (2023). A Situation Awareness Perspective on Human-AI Interaction: Tensions and Opportunities. International Journal of Human–Computer Interaction, 39(9), 1789-1806. https://doi.org/10.1080/10447318.2022.2093863

Kahneman, D. (2011). Thinking, fast and slow. macmillan.

Karran, A. J., Demazure, T., Hudon, A., Senecal, S., & Léger, P. M. (2022). Designing for Confidence: The Impact of Visualizing Artificial Intelligence Decisions. Front Neurosci,

16, 883385. https://doi.org/10.3389/fnins.2022.883385
Kelton, A., Pennington, R., & Tuttle, B. (2010). The Effects of Information Presentation Format on Judgment and Decision Making: A Review of the Information Systems Research. Journal of Information Systems, 24, 79-105. https://doi.org/10.2308/jis.2010.24.2.79
Krejtz, K., Duchowski, A. T., Niedzielska, A., Biele, C., & Krejtz, I. (2018). Eye tracking cognitive load using pupil diameter and microsaccades with fixed gaze. PLOS ONE, 13(9), e0203629. https://doi.org/10.1371/journal.pone.0203629
Krug, S. (2005). Don't Make Me Think!: A Common Sense Approach to Web Usability. In (2 ed.). New Riders.
Kuhail, M. A., Thomas, J., Alramlawi, S., Shah, S. J. H., & Thornquist, E. (2022). Interacting with a Chatbot-Based Advising System: Understanding the Effect of Chatbot Personality and User Gender on Behavior. Informatics, 9(4), 81. https://doi.org/10.3390/informatics9040081
Lee, J. D., & See, K. A. (2004). Trust in Automation: Designing for Appropriate Reliance. Human Factors: The Journal of the Human Factors and Ergonomics Society, 46(1), 50-80. https://doi.org/10.1518/hfes.46.1.50_30392
Lyell, D., Magrabi, F., & Coiera, E. (2018). The Effect of Cognitive Load and Task Complexity on Automation Bias in Electronic Prescribing. Human Factors, 60(7), 1008-1021. https://doi.org/10.1177/0018720818781224
Magezi, D. A. (2015). Linear mixed-effects models for within-participant psychology experiments: an introductory tutorial and free, graphical user interface (LMMgui) [Mini Review]. Frontiers in Psychology, 6. https://doi.org/10.3389/fpsyg.2015.00002
Marangunić, N., & Granić, A. (2015). Technology acceptance model: a literature review from 1986 to 2013. Universal Access in the Information Society, 14(1), 81-95. https://doi.org/10.1007/s10209-014-0348-1
Mazingue, C. (2023). Perceived Challenges and Benefits of AI Implementation in Customer Relationship Management Systems. Journal of Digitovation and Information System, 3(1), 72-98. https://doi.org/10.54433/jdiis.2023100023
Mejtoft, T., Hale, S., & Söderström, U. (2019). Design friction. Proceedings of the 31st European Conference on Cognitive Ergonomics,
Menekse Dalveren, G. G., & Cagiltay, N. E. (2019). Are left- and right-eye pupil sizes always equal? Journal of Eye Movement Research, 12(2). https://doi.org/10.16910/jemr.12.2.1
Merritt, S. M., Ako-Brew, A., Bryant, W. J., Staley, A., McKenna, M., Leone, A., & Shirase, L. (2019). Automation-Induced Complacency Potential: Development and Validation of a New Scale [Original Research]. Frontiers in Psychology, 10. https://doi.org/10.3389/fpsyg.2019.00225
Metcalfe, J., & Greene, M. J. (2007). Metacognition of agency. Journal of Experimental Psychology: General, 136(2), 184.
Myers, B. A. (1998). A brief history of human-computer interaction technology. interactions, 5(2), 44-54.

Naiseh, M., Al-Mansoori, R. S., Al-Thani, D., Jiang, N., & Ali, R. (2021). Nudging through friction: An approach for calibrating trust in explainable ai. 2021 8th International Conference on Behavioral and Social Computing (BESC),

Pagliari, M., Chambon, V., & Berberian, B. (2022). What is new with Artificial Intelligence? Human–agent interactions through the lens of social agency [Review]. Frontiers in Psychology, 13. https://doi.org/10.3389/fpsyg.2022.954444

Parasuraman, R., & Manzey, D. H. (2010). Complacency and Bias in Human Use of Automation: An Attentional Integration. Human Factors, 52(3), 381-410. https://doi.org/10.1177/0018720810376055

Pattyn, N., Neyt, X., Henderickx, D., & Soetens, E. (2008). Psychophysiological investigation of vigilance decrement: boredom or cognitive fatigue? Physiol Behav, 93(1-2), 369-378. https://doi.org/10.1016/j.physbeh.2007.09.016

Peng, S., Kalliamvakou, E., Cihon, P., & Demirer, M. (2023). The impact of ai on developer productivity: Evidence from github copilot. arXiv preprint arXiv:2302.06590.

Petkovic, J., Welch, V., Jacob, M. H., Yoganathan, M., Ayala, A. P., Cunningham, H., & Tugwell, P. (2016). The effectiveness of evidence summaries on health policymakers and health system managers use of evidence from systematic reviews: a systematic review. Implementation Science, 11(1). https://doi.org/10.1186/s13012-016-0530-3

Pouw, W., Rop, G., De Koning, B., & Paas, F. (2019). The cognitive basis for the split-attention effect. Journal of Experimental Psychology: General, 148(11), 2058.

Samoili, S., Cobo, M. L., Gómez, E., De Prato, G., Martínez-Plumed, F., & Delipetrev, B. (2020). AI Watch. Defining Artificial Intelligence. Towards an operational definition and taxonomy of artificial intelligence.

Searle, S. R., Speed, F. M., & Milliken, G. A. (1980). Population marginal means in the linear model: an alternative to least squares means. The American Statistician, 34(4), 216-221.

Shleifer, A. (2012). Psychologists at the gate: a review of Daniel Kahneman's thinking, fast and slow. Journal of Economic Literature, 50(4), 1080-1091.

Shneiderman, B. (2022). Human-centered AI: ensuring human control while increasing automation. Proceedings of the 5th Workshop on Human Factors in Hypertext,

Starkweather, J. (2010). Linear mixed effects modeling using R. Unpublished Manuscript.

Tobii. (2023, 01/06/2023). Measuring Pupil Size. https://connect.tobii.com/s/article/measuring-pupil-size?language=en_US

Tobii. (2024, 24/04/2024). Eye tracker sampling frequency. https://connect.tobii.com/s/article/eye-tracker-sampling-frequency?language=en_US

Wanner, J., Herm, L.-V., Heinrich, K., Janiesch, C., & Zschech, P. (2020). White, Grey, Black: Effects of XAI Augmentation on the Confidence in AI-based Decision Support Systems. ICIS,

Wen, W., Yamashita, A., & Asama, H. (2015). The sense of agency during continuous action: performance is more important than action-feedback association. PloS one, 10(4), e0125226.

Wu, X., Xiao, L., Sun, Y., Zhang, J., Ma, T., & He, L. (2022). A survey of human-in-the-loop for machine learning. Future Generation Computer Systems, 135, 364-381. https://doi.org/10.1016/j.future.2022.05.014

Xu, W. (2019). Toward human-centered AI: a perspective from human-computer interaction. interactions, 26(4), 42-46.

Xu, W., Dainoff, M. J., Ge, L., & Gao, Z. (2023). Transitioning to human interaction with AI systems: New challenges and opportunities for HCI professionals to enable human-centered AI. International Journal of Human–Computer Interaction, 39(3), 494-518.

Yetgin, E., Jensen, M., & Shaft, T. (2015). Complacency and Intentionality in IT Use and Continuance. AIS Transactions on Human-Computer Interaction, 7(1), 17-42. https://doi.org/10.17705/1thci.00064

Young, M., & Stanton, N. (2002). Attention and automation: New perspectives on mental underload and performance. Theoretical Issues in Ergonomics Science, 3. https://doi.org/10.1080/14639220210123789

Chapter 2: Enhancing Human-Computer Interaction in the Workplace

Abstract

In the context of rapid AI advancements, designing systems that support human decision-making has become critical. Our recent study explored various AI interaction modes and their effects on user perceptions, cognitive engagement, and decision-making. The findings indicate that AI systems can enhance user engagement and decision-making through deliberate 'interaction friction'. This article provides actionable recommendations for AI designers to create user-friendly systems that promote active participation, autonomy, and thoughtful decision-making, based on our research findings.

2.1. Introduction

According to a survey conducted by Microsoft and LinkedIn in 2024, at least 75% of global knowledge workers declare using Artificial Intelligence (AI) in their work (Microsoft & LinkedIn, 2024). AI's integration into job functions has led to significant improvements in efficiency and productivity. For instance, in sectors like finance, healthcare, and customer service, AI systems are used to analyse large datasets, provide insights, and support complex decision-making. These advancements aim to allow employees to focus their resources on more strategic and creative tasks.

Currently, the prevailing trend among AI designers has been to minimise user effort by ensuring that interactions between AI and humans are as smooth as possible. This approach is based on the idea that reducing cognitive load and simplifying interactions will lead to higher productivity and user satisfaction. However, the generalised integration of AI also poses challenges. As AI systems take on more responsibilities, the role and obligations of employees in their collaboration become blurrier. In the medical and transportation sectors, the challenges of automation bias and complacency become increasingly significant as users' overreliance on AI leads to disengagement and performance concerns. To address these issues, we decided to follow a new research call by integrating some intentional friction in AI designs. Interaction friction involves deliberately designing AI interactions that integrate elements encouraging users to engage more deeply with the system. This approach suggests that a certain level of friction can enhance the user experience by promoting active participation.

In our research, participants were required to perform utilitarian tasks, simulating the experience of employees filling out forms on their company's intranet. We tested two specific examples of interaction friction in chatbots. The summary feature provides users with a summary of tasks performed with the AI, encouraging users to verify the outputs provided by the AI. The form-based chatbot incorporates a form that the user must fill out alongside interacting with the AI chatbot, requiring users to read and verify the information they provide. The figure below illustrates the three modes of AI interaction evaluated in our study: the first depicts the standard, frictionless chatbot interface, while the subsequent screenshots show the two friction-inducing designs, the middle one with the summary and the right with the form-based chatbot.

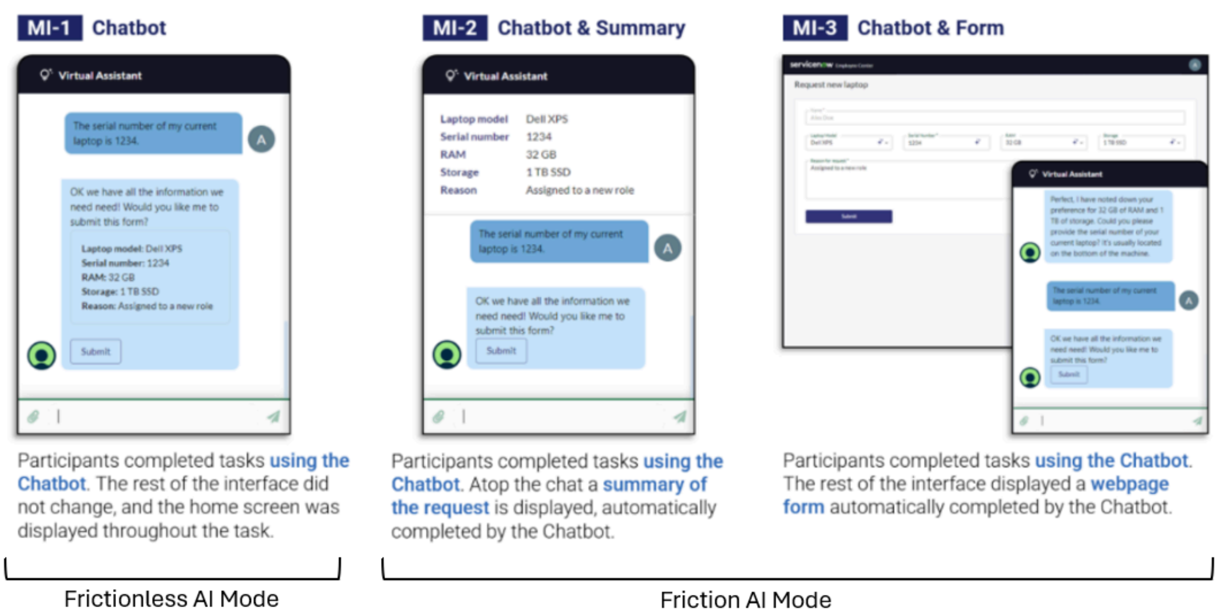

Figure 4: Different modes of chatbots

By incorporating these examples of design friction, our study aims to demonstrate how thoughtful design can enhance users' perceptions, cognitive states, and task performance. The following sections will present the benefits observed from adding friction to AI interaction in our study.

2.2. Enhancing Perceived Usefulness

In our study, one of the key findings was that users found the system most useful when the summary feature was available. This was in comparison to both the no-AI mode and the chatbot with a form. The summary feature provided users with an overview of the tasks performed with the AI, allowing them to easily verify and understand the outputs generated. For instance, imagine an employee requesting a new work laptop through an AI system. The summary feature would allow the employee to review all the details of their request, such as verifying that the laptop meets their specific prerequisites and confirming the correct delivery location. This summary helps to ensure that the employee's request is aligned with their needs, reducing the likelihood of errors.

Perceived usefulness is a crucial factor in user satisfaction and intention to use AI systems. When users find an AI system useful, they are more likely to have a positive opinion of the AI system, leading to higher levels of satisfaction and continued use. To illustrate this point, consider a specific example from our research. In the comparison between the chatbot with the summary mode and the no-AI condition, there was a highly significant increase in perceived usefulness when the summary feature was available. This finding indicates that users found the friction introduced by the summary feature more beneficial than performing the task manually. Furthermore, when comparing the chatbot with the summary to the chatbot with the form, the presence of a summary was perceived as more useful than the form-based chatbot design. In contrast, performing the task frictionlessly with a regular chatbot did not have any significant effect on perceived usefulness compared to performing the task manually.

This evidence suggests that adding a certain level of friction, such as providing a summary of AI actions, can significantly enhance the perceived usefulness of AI systems. It highlights the importance of thoughtful design in AI interactions, where the right balance of friction can lead to better perceptions of the AI assistants.

2.3. Building Confidence and Trust

One of the critical findings from our study was that users reported higher confidence in their outputs and greater trust in the system when AI was present, regardless of the AI interaction mode. This result was observed across various AI interaction modes, including the frictionless chatbot-only mode, the chatbot with a summary mode, and the chatbot with a form mode. In each case, users exhibited significantly higher confidence compared to the mode where no AI was used. This suggests that the presence of AI itself, rather than the specific interaction mode, is an important factor in enhancing user confidence in task completion. For example, in our study, participants were required to request their pay stubs through the system. When AI was involved in the process, users reported feeling more confident in the accuracy of their request. This was particularly important for such sensitive information, as the AI system was responsible for validating and correcting the inputs provided by the employees. The reassurance that the AI was double-checking their entries significantly boosted users' confidence in the correctness of their forms, reducing the anxiety associated with potential errors.

Confidence and trust are foundational to successful user interactions with AI systems. Higher confidence in the outputs provided by AI can lead to better decision-making, as users are more likely to rely on and act upon the information given by the AI. Trust in the AI system is equally important, as it is one of the antecedents of users' intention to use the system. Without trust, users may be hesitant to use AI systems, denying the potential benefits these technologies can offer. From a managerial standpoint, this highlights the importance of incorporating AI in tasks where accuracy and trust are critical, such as payroll or other sensitive employee data management. By doing so, organisations can enhance employees' confidence in their tasks, leading to fewer errors and more efficient processes. Furthermore, building trust in AI systems can increase adoption rates, ensuring that employees feel comfortable and supported when using these technologies.

These findings underscore the importance of incorporating AI into user interactions to boost both confidence and trust, which are essential for effective and satisfying user experiences.

2.4. Evaluating Performance

One of the key findings of our study was that no significant difference was found in the time taken to complete tasks between the friction AI modes and non-AI conditions. Specifically, the results showed that while the time to complete tasks was longer when participants used the chatbot-only mode compared to performing the task manually without AI assistance, there were no significant differences in task performance time between the other Friction AI modes (chatbot with summary and chatbot with form) and the control mode. This suggests that adding certain types of interaction friction does not slow down task performance. Through various scenarios, the addition of friction, such as the summary or form modes, didn’t slow down participants but rather encouraged them to verify the outputs and requests they sent to the AI. These extra features in the interaction, combined with the AI’s ability to enter and validate the participant's inputs, means that while the speed of task completion was not affected, the potential quality of the output was improved. Both the users and the AI were engaged in verifying the requests, leading to more accurate and reliable outcomes.

These results suggest that the thoughtful design of interaction friction can enhance user perceptions and satisfaction without compromising efficiency. By carefully adding some frictions in AI, designers can ensure that users remain involved without experiencing delays in task completion. This balance is important for developing effective and user-friendly AI systems that both perform well and meet users' needs. From a managerial perspective, this finding demonstrates that efficiency can be maintained even as measures are put in place to improve the quality and reliability of outputs. By encouraging users to engage more thoroughly with the AI system, organisations can achieve higher-quality results without sacrificing speed.

2.5. Conclusion

The findings indicate that AI systems, when designed with varying levels of interaction friction, can greatly impact users' perceptions of the AI assistant's usefulness and trustworthiness. Our study showed that by incorporating deliberate interaction friction, AI systems can enhance users' perceptions of system usefulness, confidence, and trust. In addition, it was shown that adding friction does not affect users’ task performance, as no difference was observed between friction AI modes and manual task times. For example, in software design, designers could integrate multiple positive frictions such as a prompt in the chatbot to request users' revision before submitting, or presenting users with explanations for the AI decisions. These results highlight the potential for thoughtful AI design to improve user experiences without compromising efficiency.

Methodology

The recommendations presented in this article are derived from a research project that studied the effects of different AI interaction modes on user perceptions, cognitive engagement, and decision-making. 33 participants from a Canadian business school were invited. We combined data from questionnaires measuring perceived usefulness, confidence, trust, and sense of agency with physiological measurements, such as eye tracking, to assess cognitive load of the participants.

Reference list Article 2

Microsoft, & LinkedIn. (2024). 2024 Work Trend Index Annual Report. https://www.microsoft.com/en-us/worklab/work-trend-index/ai-at-work-is-here-now-comes-the-hard-part

Conclusion

Research Objectives and Methodology

This thesis aimed to explore the impact of interaction friction in AI-human interactions to enhance user experience and decision-making processes. The research was conducted through a lab experiment where the participants interacted with a Generative AI Assistant (GENIAI) while their perceptions, emotional and cognitive states were recorded. The study combined quantitative data, observational data, and physiological measurements. This methodology allowed for a nuanced analysis of user perceptions and cognitive responses.

The first article, titled "Interaction Friction in AI: Improving User Experience and Decision Making," focused on understanding the effects of various AI interaction modes on user's perception and cognitive states. This project studied how different levels and types of interaction friction influence users' engagement with AI systems and their cognitive processing during tasks. By introducing controlled friction, such as prompts and windows for users' feedback, the study aimed to understand how these elements could potentially lead to better decision-making outcomes and heightened user perception.

The second article, "Enhancing Human-Computer Interaction in the Workplace," expanded on these concepts by presenting insights on how different designs of AI frictions can be applied in workplace settings to improve user experience and performance. This article was focused on practical applications, examining how deliberate interaction friction could enhance user satisfaction, confidence, trust and sense of agency in professional environments. It aimed to provide insights into the implementation of friction elements in AI systems used in the workplace, therefore bridging the gap between theoretical research and real-world applications.

Revisiting Research Questions and Hypotheses

The primary research question guiding this project was: To what extent do various AI friction affect user perceptions, cognitive state, and behaviour? This question aimed to explore the dynamics between users and AI systems, focusing on how different interaction designs can influence user experience and perception of AI applications. The goal was to uncover the specific ways in which varying levels of interaction friction impact users' cognitive processes and their overall decision-making abilities.

To address this research question, several hypotheses were formulated and tested. The outcomes of these hypotheses provide insights into the impact of various AI interaction modes on user perceptions, cognitive states, and decision-making processes.

The study found that various AI interaction modes impact user perceptions, cognitive states, and decision-making processes in different ways. Controlled interaction friction, such as summaries and form features, improved the perceived usefulness of the AI systems. However, cognitive load, confidence, trust in AI, and sense of agency did not significantly decrease between different friction modes and frictionless modes of AI. It was the presence of AI itself, rather than the friction elements, that increased users' confidence, trust, and perceived sense of agency while also reducing cognitive load. Importantly, the introduction of friction did not significantly decrease these perceptions, nor did the design of the friction elements significantly increase cognitive load or task completion time. Overall, the findings suggest that controlled interaction friction can enhance user experience, cognitive state and decision-making in AI interaction.

Theoretical Contributions

This research makes several theoretical contributions to the fields of Human-Computer Interaction (HCI) and AI, particularly in understanding how interaction design can influence user experience and decision-making processes.

One of the primary contributions is the exploration of the interaction friction concept. This idea challenges the traditional HCI approach that has primarily focused on creating frictionless interactions so as to maximise efficiency and user satisfaction. Traditionally, HCI has encouraged reducing user effort to make interactions with technology as effortless as possible. However, this research demonstrates that deliberately introducing controlled friction can have positive effects on users’ perceptions. The interaction friction concept provides a new perspective for designing human AI interactions, suggesting that a certain amount of friction can enhance the quality of user experiences by encouraging the user to stay more engaged in the task.

The study conducted in this research highlights the importance of balancing user perceptions and cognitive load. While traditional HCI approaches have aimed to minimise cognitive load, this research indicates that controlled interaction friction can reduce cognitive effort while still enhancing user experience. Specifically, the findings show that users experienced a lower cognitive load when using AI assistance compared to performing tasks manually. Despite making less cognitive effort with AI assistance, users reported feeling more in control (a higher sense of agency), and declared having increased confidence and trust in the AI systems. By introducing elements such as summaries and prompt for feedback, the AI systems encouraged deeper engagement with the tasks, leading to improved decision-making processes. These results suggest that a balance can be achieved where cognitive load is reduced, yet users remain actively engaged and empowered, highlighting the potential benefits of controlled interaction friction in enhancing user experience.

Practical Implications

The practical implications of this study are twofold. On one hand, they provide guidance for AI designers on how intentional frictions can enhance users' perceptions. On the other hand, they encourage a more thoughtful approach to human-AI collaboration in the workplace.

One key implication is that designers would benefit from incorporating controlled interaction friction into AI systems. By integrating elements that require user input or decision points, designers can increase user engagement and improve decision-making processes. For example, AI systems can include summaries of the tasks or prompts to ask for users feedback. These interaction points keep users actively involved, encouraging a sense of responsibility over the task and its outcomes. Shifting from a frictionless design to one that strategically uses friction can improve the user experience with the systems.

In workplace settings, applying interaction friction can improve employees performance and perceptions. AI systems with intentionally designed friction promote active participation and responsibility among employees. For instance, decision-assistant systems can include prompts for user confirmation or detailed feedback requests in the form of summaries to enhance decision quality. This approach helps reduce automation complacency, where users might rely too heavily on AI without critical evaluation. By encouraging more engaged and thoughtful interactions, AI systems can help employees make better decisions and stay cognitively involved in their tasks.

Appendices

Construct	Measurement	Measurement tool	Reference
Perceived Usefulness	Scale (1-7)	Using this system in my job would enable me to accomplish tasks more quickly. Using this system would improve my job performance. Using this system in my job would increase my productivity. Using this system would enhance my effectiveness on the job. Using this system would make it easier to do my job.	(Davis, 1989)
Sense of Agency	Scale (1-100)	How much in control were you during this task?	(Metcalfe & Greene, 2007)
Confidence in AI	Scale (1-7)	I am convinced of the output given to me by the system.	(Falconnet et al., 2023)
Trust in AI	Scale (1-7)	The system can be trusted.	(Falconnet et al., 2023)
Cognitive Load	Pupil dilatation (mm)	Eye Tracker Tobii Pro Lab (Tobii AB, Stockholm, Sweden)	(Beatty, 1982)

Reference list

Abdi, H., & Williams, L. J. (2010). Tukey's honestly significant difference (HSD) test. Encyclopedia of research design, 3(1), 1-5.

Adamopoulou, E., & Moussiades, L. (2020). Chatbots: History, technology, and applications. Machine Learning with Applications, 2, 100006. https://doi.org/https://doi.org/10.1016/j.mlwa.2020.100006

Amershi, S., Weld, D., Vorvoreanu, M., Fourney, A., Nushi, B., Collisson, P.,…Horvitz, E. (2019, 2019). Guidelines for Human-AI Interaction.

Baecker, R. M. (2008). TIMELINES Themes in the early history of HCI---some unanswered questions. Interactions, 15(2), 22-27.

Beatty, J. (1982). Task-evoked pupillary responses, processing load, and the structure of processing resources. Psychological Bulletin, 91(2), 276-292. https://doi.org/10.1037/0033-2909.91.2.276

Brachten, F., Brünker, F., Frick, N. R. J., Ross, B., & Stieglitz, S. (2020). On the ability of virtual agents to decrease cognitive load: an experimental study. Information Systems and e-Business Management, 18(2), 187-207. https://doi.org/10.1007/s10257-020-00471-7

Buckley, R. P., Zetzsche, D. A., Arner, D. W., & Tang, B. W. (2021). Regulating artificial intelligence in finance: Putting the human in the loop. Sydney Law Review, The, 43(1), 43-81.

Cambridge University Press. (2024). Friction. In Cambridge Academic Content Dictionary. https://dictionary.cambridge.org/dictionary/english/friction

Casadei, A., Schlogl, S., & Bergmann, M. (2022). Chatbots for Robotic Process Automation: Investigating Perceived Trust and User Satisfaction.

Chen, Z., & Schmidt, R. (2024a). Exploring a Behavioral Model of "Positive Friction" in Human-AI Interaction. International Conference on Human-Computer Interaction, Chen, Z., & Schmidt, R. (2024b). Exploring a Behavioral Model of "Positive Friction" in Human-AI Interaction. In (pp. 3-22). Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-61353-1_1

Cheng, Y., & Jiang, H. (2020). How Do AI-driven Chatbots Impact User Experience? Examining Gratifications, Perceived Privacy Risk, Satisfaction, Loyalty, and Continued Use. Journal of Broadcasting & Electronic Media, 64(4), 592-614. https://doi.org/10.1080/08838151.2020.1834296

Chong, L., Zhang, G., Goucher-Lambert, K., Kotovsky, K., & Cagan, J. (2022). Human confidence in artificial intelligence and in themselves: The evolution and impact of confidence on adoption of AI advice. Computers in Human Behavior, 127, 107018. https://doi.org/https://doi.org/10.1016/j.chb.2021.107018

Chong, T., Yu, T., Keeling, D. I., & de Ruyter, K. (2021). AI-chatbots on the services frontline addressing the challenges and opportunities of agency. Journal of Retailing and Consumer Services, 63. https://doi.org/10.1016/j.jretconser.2021.102735

Cox, A. L., Gould, S. J., Cecchinato, M. E., Iacovides, I., & Renfree, I. (2016). Design frictions for mindful interactions: The case for microboundaries. Proceedings of the 2016 CHI conference extended abstracts on human factors in computing systems,

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

De Bruyne, J., Joundi, J., Morton, J., Zheleva, A., Van Kets, N., Van Wallendael, G.,…Bombeke, K. (2023). I spy with my AI: The effects of AI-based visual cueing on human operators' performance and cognitive load in CCTV control rooms. International Journal of Industrial Ergonomics, 95, 103444. https://doi.org/https://doi.org/10.1016/j.ergon.2023.103444

Endsley, M. R. (1995). Toward a Theory of Situation Awareness in Dynamic Systems. Human Factors: The Journal of the Human Factors and Ergonomics Society, 37(1). https://doi.org/info:doi/10.1518/001872095779049543

Falconnet, A., Coursaris, C. K., Beringer, J., Van Osch, W., Sénécal, S., & Léger, P.-M. (2023). Improving User Experience with Recommender Systems by Informing the Design of

Recommendation Messages. Applied Sciences, 13(4), 2706. https://doi.org/10.3390/app13042706

Fox, J. R., Park, B., & Lang, A. (2007). When Available Resources Become Negative Resources:The Effects of Cognitive Overload on Memory Sensitivity and Criterion Bias. Communication Research, 34(3), 277-296. https://doi.org/10.1177/0093650207300429

Friedland, J. (2020). AI can help us live more deliberately.

Fujita, H. (2020). AI-based computer-aided diagnosis (AI-CAD): the latest review to read first. Radiological Physics and Technology, 13(1), 6-19. https://doi.org/10.1007/s12194-019-00552-4

Gkinko, L., & Elbanna, A. (2022). Hope, tolerance and empathy: employees' emotions when using an AI-enabled chatbot in a digitalised workplace. Information Technology & People, 35(6), 1714-1743. https://doi.org/10.1108/itp-04-2021-0328

Goodell, J. W., Kumar, S., Lim, W. M., & Pattnaik, D. (2021). Artificial intelligence and machine learning in finance: Identifying foundations, themes, and research clusters from bibliometric analysis. Journal of Behavioral and Experimental Finance, 32, 100577. https://doi.org/https://doi.org/10.1016/j.jbef.2021.100577

Grudin, J. (2008). A moving target: The evolution of HCI. The human-computer interaction handbook: Fundamentals, evolving technologies, and emerging applications, 1-24.

Gönül, M. S., Önkal, D., & Lawrence, M. (2006). The effects of structural characteristics of explanations on use of a DSS. Decision support systems, 42(3), 1481-1493.

Haan, K. (2023). How Businesses Are Using Artificial Intelligence In 2024. Forbes Advisor. https://www.forbes.com/advisor/business/software/ai-in-business/

Hemmer, P., Westphal, M., Schemmer, M., Vetter, S., Vössing, M., & Satzger, G. (2023, 2023). Human-AI Collaboration: The Effect of AI Delegation on Human Task Performance and Task Satisfaction.

Holmes, W., & Tuomi, I. (2022). State of the art and practice in AI in education. European Journal of Education, 57(4), 542-570.

Hudon, A., Demazure, T., Karran, A., Léger, P.-M., & Sénécal, S. (2021). Explainable Artificial Intelligence (XAI): How the Visualization of AI Predictions Affects User Cognitive Load and Confidence. In (pp. 237-246). Springer International Publishing. https://doi.org/10.1007/978-3-030-88900-5_27

Hwang, S., & Kim, J. (2021). Toward a Chatbot for Financial Sustainability. Sustainability, 13(6), 3173.

Jiang, F., Jiang, Y., Zhi, H., Dong, Y., Li, H., Ma, S.,…Wang, Y. (2017). Artificial intelligence in healthcare: past, present and future. Stroke and vascular neurology, 2(4).

Jiang, J., Karran, A. J., Coursaris, C. K., Léger, P.-M., & Beringer, J. (2023). A Situation Awareness Perspective on Human-AI Interaction: Tensions and Opportunities. International Journal of Human–Computer Interaction, 39(9), 1789-1806. https://doi.org/10.1080/10447318.2022.2093863

Kahneman, D. (2011). Thinking, fast and slow. macmillan.

Karran, A. J., Demazure, T., Hudon, A., Senecal, S., & Léger, P. M. (2022). Designing for Confidence: The Impact of Visualizing Artificial Intelligence Decisions. Front Neurosci, 16, 883385. https://doi.org/10.3389/fnins.2022.883385

Kelton, A., Pennington, R., & Tuttle, B. (2010). The Effects of Information Presentation Format on Judgment and Decision Making: A Review of the Information Systems Research. Journal of Information Systems, 24, 79-105. https://doi.org/10.2308/jis.2010.24.2.79

Krejtz, K., Duchowski, A. T., Niedzielska, A., Biele, C., & Krejtz, I. (2018). Eye tracking cognitive load using pupil diameter and microsaccades with fixed gaze. PLOS ONE, 13(9), e0203629. https://doi.org/10.1371/journal.pone.0203629

Krug, S. (2005). Don't Make Me Think!: A Common Sense Approach to Web Usability. In (2 ed.). New Riders.

Kuhail, M. A., Thomas, J., Alramlawi, S., Shah, S. J. H., & Thornquist, E. (2022). Interacting with a Chatbot-Based Advising System: Understanding the Effect of Chatbot Personality and User Gender on Behavior. Informatics, 9(4), 81. https://doi.org/10.3390/informatics9040081

Lee, J. D., & See, K. A. (2004). Trust in Automation: Designing for Appropriate Reliance. Human Factors: The Journal of the Human Factors and Ergonomics Society, 46(1), 50-80. https://doi.org/10.1518/hfes.46.1.50_30392

Lyell, D., Magrabi, F., & Coiera, E. (2018). The Effect of Cognitive Load and Task Complexity on Automation Bias in Electronic Prescribing. Human Factors, 60(7), 1008-1021. https://doi.org/10.1177/0018720818781224

Magezi, D. A. (2015). Linear mixed-effects models for within-participant psychology experiments: an introductory tutorial and free, graphical user interface (LMMgui) [Mini Review]. Frontiers in Psychology, 6. https://doi.org/10.3389/fpsyg.2015.00002

Marangunić, N., & Granić, A. (2015). Technology acceptance model: a literature review from 1986 to 2013. Universal Access in the Information Society, 14(1), 81-95. https://doi.org/10.1007/s10209-014-0348-1

Mazingue, C. (2023). Perceived Challenges and Benefits of AI Implementation in Customer Relationship Management Systems. Journal of Digitovation and Information System, 3(1), 72-98. https://doi.org/10.54433/jdiis.2023100023

Mejtoft, T., Hale, S., & Söderström, U. (2019). Design friction. Proceedings of the 31st European Conference on Cognitive Ergonomics,

Menekse Dalveren, G. G., & Cagiltay, N. E. (2019). Are left- and right-eye pupil sizes always equal? Journal of Eye Movement Research, 12(2). https://doi.org/10.16910/jemr.12.2.1

Merritt, S. M., Ako-Brew, A., Bryant, W. J., Staley, A., McKenna, M., Leone, A., & Shirase, L. (2019). Automation-Induced Complacency Potential: Development and Validation of a New Scale [Original Research]. Frontiers in Psychology, 10. https://doi.org/10.3389/fpsyg.2019.00225

Messeri, C., Masotti, G., Zanchettin, A. M., & Rocco, P. (2021). Human-Robot Collaboration: Optimizing Stress and Productivity Based on Game Theory. IEEE Robotics and Automation Letters, 6(4), 8061-8068. https://doi.org/10.1109/LRA.2021.3102309

Metcalfe, J., & Greene, M. J. (2007). Metacognition of agency. Journal of Experimental Psychology: General, 136(2), 184.

Microsoft, & LinkedIn. (2024). 2024 Work Trend Index Annual Report. https://www.microsoft.com/en-us/worklab/work-trend-index/ai-at-work-is-here-now-comes-the-hard-part

Myers, B. A. (1998). A brief history of human-computer interaction technology. interactions, 5(2), 44-54.

Naiseh, M., Al-Mansoori, R. S., Al-Thani, D., Jiang, N., & Ali, R. (2021). Nudging through friction: An approach for calibrating trust in explainable ai. 2021 8th International Conference on Behavioral and Social Computing (BESC),

Natali, C. (2023). Per Aspera ad Astra, or Flourishing via Friction: Stimulating Cognitive Activation by Design through Frictional Decision Support Systems. CEUR workshop proceedings,

Pagliari, M., Chambon, V., & Berberian, B. (2022). What is new with Artificial Intelligence? Human–agent interactions through the lens of social agency [Review]. Frontiers in Psychology, 13. https://doi.org/10.3389/fpsyg.2022.954444

Parasuraman, R., & Manzey, D. H. (2010). Complacency and Bias in Human Use of Automation: An Attentional Integration. Human Factors, 52(3), 381-410. https://doi.org/10.1177/0018720810376055

Pattyn, N., Neyt, X., Henderickx, D., & Soetens, E. (2008). Psychophysiological investigation of vigilance decrement: boredom or cognitive fatigue? Physiol Behav, 93(1-2), 369-378. https://doi.org/10.1016/j.physbeh.2007.09.016

Peng, S., Kalliamvakou, E., Cihon, P., & Demirer, M. (2023). The impact of ai on developer productivity: Evidence from github copilot. arXiv preprint arXiv:2302.06590.

Petkovic, J., Welch, V., Jacob, M. H., Yoganathan, M., Ayala, A. P., Cunningham, H., & Tugwell, P. (2016). The effectiveness of evidence summaries on health policymakers and health system managers use of evidence from systematic reviews: a systematic review. Implementation Science, 11(1). https://doi.org/10.1186/s13012-016-0530-3

Pouw, W., Rop, G., De Koning, B., & Paas, F. (2019). The cognitive basis for the split-attention effect. Journal of Experimental Psychology: General, 148(11), 2058.

Roto, V., Palanque, P., & Karvonen, H. (2019). Engaging Automation at Work – A Literature Review. In (pp. 158-172). Springer International Publishing. https://doi.org/10.1007/978-3-030-05297-3_11

Samoili, S., Cobo, M. L., Gómez, E., De Prato, G., Martínez-Plumed, F., & Delipetrev, B. (2020). AI Watch. Defining Artificial Intelligence. Towards an operational definition and taxonomy of artificial intelligence.

Schmager, S., Pappas, I., & Vassilakopoulou, P. (2023). Defining human-centered AI: a comprehensive review of HCAI literature.

Searle, S. R., Speed, F. M., & Milliken, G. A. (1980). Population marginal means in the linear model: an alternative to least squares means. The American Statistician, 34(4), 216-221.

Shleifer, A. (2012). Psychologists at the gate: a review of Daniel Kahneman's thinking, fast and slow. Journal of Economic Literature, 50(4), 1080-1091.

Shneiderman, B. (2022). Human-Centered AI. OUP Oxford.

Shneiderman, B. (2022). Human-centered AI: ensuring human control while increasing automation. Proceedings of the 5th Workshop on Human Factors in Hypertext,

Starkweather, J. (2010). Linear mixed effects modeling using R. Unpublished Manuscript.

Sun, Y., Li, S., & Yu, L. (2022). The dark sides of AI personal assistant: effects of service failure on user continuance intention. Electronic Markets, 32(1), 17-39.

Tobii. (2023, 01/06/2023). Measuring Pupil Size. https://connect.tobii.com/s/article/measuring-pupil-size?language=en_US

Tobii. (2024, 24/04/2024). Eye tracker sampling frequency. https://connect.tobii.com/s/article/eye-tracker-sampling-frequency?language=en_US

Wanner, J., Herm, L.-V., Heinrich, K., Janiesch, C., & Zschech, P. (2020). White, Grey, Black: Effects of XAI Augmentation on the Confidence in AI-based Decision Support Systems. ICIS,

Wen, W., Yamashita, A., & Asama, H. (2015). The sense of agency during continuous action: performance is more important than action-feedback association. PloS one, 10(4), e0125226.

Wu, X., Xiao, L., Sun, Y., Zhang, J., Ma, T., & He, L. (2022). A survey of human-in-the-loop for machine learning. Future Generation Computer Systems, 135, 364-381. https://doi.org/10.1016/j.future.2022.05.014

Xu, W. (2019). Toward human-centered AI: a perspective from human-computer interaction. interactions, 26(4), 42-46.

Xu, W., Dainoff, M. J., Ge, L., & Gao, Z. (2023). Transitioning to human interaction with AI systems: New challenges and opportunities for HCI professionals to enable human-centered AI. International Journal of Human–Computer Interaction, 39(3), 494-518.

Yetgin, E., Jensen, M., & Shaft, T. (2015). Complacency and Intentionality in IT Use and Continuance. AIS Transactions on Human-Computer Interaction, 7(1), 17-42. https://doi.org/10.17705/1thci.00064

Young, M., & Stanton, N. (2002). Attention and automation: New perspectives on mental underload and performance. Theoretical Issues in Ergonomics Science, 3. https://doi.org/10.1080/14639220210123789

Zhang, K., & Aslan, A. B. (2021). AI technologies for education: Recent research & future directions. Computers and Education: Artificial Intelligence, 2, 100025. https://doi.org/https://doi.org/10.1016/j.caeai.2021.100025