AI vs. Humans: Who do we trust?

Examining trust in AI among Swedish adults

REBECKA BERG
GUSTAV LINDER

Bachelor Thesis

Stockholm School of Economics

2023

AI vs. Humans: Who do we trust? A survey study examining trust in AI in Swedish adults.

Abstract:

This study investigates trust in Artificial Intelligence (AI) among Swedish adults, focusing on how different factors influence this trust and the nuances of AI-human decision-making. A quantitative survey, distributed mainly on Swedish university campuses, centered around four scenarios: Dating, Math, Doctor and Career, aimed at gauging emotional and cognitive trust in AI and the preference for AI or human recommendations across various contexts. Analysis of responses from 195 participants brought to light several key insights. Firstly, cognitive trust in AI tended to surpass emotional trust and this varied depending on the subjectivity of the scenarios presented. Secondly, while participants often showed greater trust in human recommendations, this tendency fluctuate based on how objective they perceived the task to be. Furthermore, there was a positive correlation between experience with AI and trust, though this was not consistent across different scenarios. Lastly, a notable negative correlation was observed between technological anxiety and trust in AI. The study concludes that task nature significantly affects trust in AI, particularly in tasks seen as objective, offering important insights for AI's role in marketing and policy-making. The research recognizes demographic limitations and suggests future exploration into AI trust across diverse populations and evolving AI applications.

Keywords:

Trust in AI, AI-human interaction, Cognitive trust, Emotional trust, Decision-making, Subjectivity in AI, Technology anxiety, AI in marketing, Task objectivity

Authors:

Rebecka Berg (25331)
Gustav Linder (25285)

Tutors:

Patric Andersson, Associate Professor, Department of Marketing and Strategy

Examiner: SSE faculty

Bachelor Thesis

Bachelor Program in Business and Economics
Stockholm School of Economics

Table of Contents

Acknowledgements

We would like to thank everyone who helped us in completing this thesis. We extend our special thanks and deep appreciation to:

Patric Andersson

For being an excellent support at all times and providing valuable knowledge, knowledge, and guidance.

All respondents

Without you this thesis would not have been possible. Thank you for taking the time and involvement by answering our survey.

Friends and family

For the support and always being there.

Yours sincerely,
Rebecka and Gustav

Definitions

Artificial Intelligence (AI): Artificial Intelligence (AI) is the ability of computer programs/robots to mimic human natural intelligence. This includes the capability to learn from previous experiences, understand natural language, solve problems, plan a sequence of actions and generalize.

Cognitive Trust in AI: Rational trust based on AI's competence, consistency and reliability (Glikson & Wolley, 2020).

Emotional Trust in AI: Trust derived from personal comfort and security with AI, beyond its technical abilities (Glikson & Woolley, 2020).

Technological Anxiety (TA): Discomfort or fear associated with using new technology (Meuter et al., 2003).

Task Characteristics: Features defining a task's nature, especially its objectivity or subjectivity (Castelo et al., 2019).

Subjectivity in AI: AI decisions influenced by personal feelings or opinions (Glikson & Woolley, 2020).

Objectivity in AI: AI decisions based on quantifiable facts, uninfluenced by personal feelings (Castelo et al., 2019).

Familiarity with AI: The frequency and extent of an individual's practical interactions with AI technologies (Glikson & Woolley, 2020).

Between-Subject Study: Research design comparing different groups under varying conditions (Bell, 2022).

1. Introduction

This section of the paper discusses the background to the topic of the thesis, namely trust in Artificial Intelligence (hereafter named AI). We will delve into the definition of AI, technological advancements, beliefs about AI, ethical aspects and important happenings in recent years. The purpose of the research is also presented.

1.1. Background

Humanity is in the midst of a transformative era: the Fourth Industrial Revolution, with AI at its core, as described by Klaus Schwab, founder and executive chairman of the World Economic Forum (Groumpos, 2021; Schwab, 2017). AI is the buzzword on everyone's lips, dominating conversations across society. It is expected to drastically alter the economy and permeate all aspects of life and business (Glikson & Woolley, 2020). Additionally, the market is projected to grow twentyfold by 2030, reaching an estimated two trillion USD. To put this in perspective, the 2021 global automotive manufacturing market was 2.56 trillion USD (Statista, 2023). The increasing presence of AI applications is highlighted by the engagement of major tech giants like Microsoft, IBM, Google and Samsung who are investing significantly in AI research and development, driving innovation as shown by their numerous AI patent applications (Statista, 2023).

Technological advancements in AI have revolutionized human-technology interaction, moving from a concept many knew of but did not fully understand to an integral part of daily life. AI has even been announced as "the word of the year" by the dictionary publisher Collins, further confirming the rapid integration of AI into individuals' lives (Collins et al., 2021). Understanding public perceptions and attitudes toward AI in different contexts is therefore vital, hence this thesis compares decision-making between humans and AI to offer insightful conclusions.

1.1.1. Industrial revolutions

We will briefly explore past revolutions to draw parallels between the reactions they evoked and the experiences we are encountering today. The First Industrial Revolution, starting in the late 18th century, marked a historic shift with steam power and mechanization, introducing innovations like the Spinning Jenny and railway locomotives (Groumpos, 2021). The Second Industrial Revolution, in the late 19th and early 20th centuries, fueled by the discovery of electricity, marked the start of mass production, one example being Henry Ford who completely changed the automobile production process (Donovan, 1997). The Third Industrial Revolution began in the late 20th century, with the advent of digital technology, internet and enabling full production automation (Groumpos, 2021). The ongoing Fourth Industrial Revolution has been unfolding since the 2000s, merging physical, digital and biological systems, primarily through AI (Xu et al., 2018).

According to some, we are entering a "jobless future" (Ford, 2015) and a "Race Against the Machine" (Brynjolfsson & Mcafee, 2011). However, history indicates that past revolutions generated job opportunities that were not foreseen at the beginning of them. This suggests

that while the future remains uncertain, historical trends demonstrate that technological advancements ultimately create new job prospects, which offers a more optimistic view of the potential job landscape of the future.

1.1.2. Definition of AI

Given the diverse definitions of AI, it is crucial to clarify what we mean by AI in this thesis. For starters, the term AI was coined in 1956 by John McCarthy, regarded as AI‘s ”Founding Father”, at a Dartmouth College conference, marking the ”birth of artificial intelligence” (Collins et al., 2021; Russell & Norvig 1995). Since then, the definition of AI has been debated as defining it is complex given its breadth and rapid evolution (World Intellectual Property Organization, 2019). The difficulty in defining AI is understandable, as many scientific concepts are defined only after maturing (Collins et al., 2021).

This thesis focuses on Narrow AI (ANI), the prevalent form of AI (Raj & Seamans, 2019; Russell & Norvig, 1995) examples include (but are not limited to) ChatGPT, Apple‘s Siri and recommendation engines on streaming/social media platforms.[1] At its core, AI is fundamentally based on machine learning (Janiesch et al., 2021). Unlike traditional automation, machine learning can learn and adapt from experience and feedback, like humans. Simply put, AI can be described as a machine‘s capability to imitate human intelligence traits like learning, creativity, reasoning and problem-solving while aiming for rational outcomes (Ertel et al., 2018; Glikson & Woolley, 2020; Russell & Norvig, 1995).

The following definition was provided in the survey (translated from Swedish): ”*Artificial Intelligence (AI) is the ability of computer programs/robots to mimic human natural intelligence. This includes the capability to learn from previous experiences, understand natural language, solve problems, plan a sequence of actions and generalize*.”

1.1.3. Ethical concerns regarding AI

“Is AI dangerous?” remains one of the most frequent searches on Google about AI (SvD, 2023). Moreover, a study by Svenska Internetstiftelsen reveals that 30% of Swedes see AI‘s future impact as predominantly negative (Svenska Internetstiftelsen, 2023). This perception may stem from media often depicting AI in a dystopian light, as seen in movies like *The Terminator* (1984), *I Robot* (2004) and *Lucy* (2014). However, with ChatGPT‘s release in November 2022 and other generative AI tools which have followed, AI‘s role in daily life has become more apparent and perhaps shifted the fear of robotic AI’s to other forms more related to legal and ethical concerns about transparency, bias and potential misuse. AI presents opportunities and risks, both of which this thesis aims to acknowledge.

AI offers opportunities like enhanced hiring efficiency by screening CVs and motivational letters, yet it also has drawbacks. For instance, AI can inherit biases from its training data, as seen with Amazon‘s AI hiring model favoring male candidates due to historical gender

[1] *AI can further be divided into subgroups such as Artificial General Intelligence (AGI, excelling human capabilities in all aspects). As AGI does not yet exist (Glikson & Woolley, 2020), it is not the focus of this thesis.*

imbalances (Forbes, 2023). Other examples include the rise of deepfakes used for malicious purposes (such as spreading fake news or committing identity theft) and copyright issues, one well known example being a hit song released in April 2023 were the artists Drake and the Weeknd‘s voices were replicated by an AI (Harvard Law Today, 2023). Another notable event that got attention worldwide was the firing of Sam Altman from OpenAI in November 2023, due to concerns about deviating from founding principles and irresponsibly advancing super intelligent AI (New York Post, 2023). Recently, there has emerged a new phenomenon where young men are opting for AI girlfriends over real-life relationships, a trend that raises significant ethical concerns about the impact of intimate chatbots on human connections and loneliness (P3 Nyheter, 2023). Lastly, AI‘s potential in military and government use, such as AI-driven malware or misuse of autonomous weapons systems, presents risks like cybersecurity breaches and military incidents (Blauth et al., 2022).

All these situations underscore the importance of ethical considerations, transparency and legal frameworks for the development and utilization of AI. However, effectively regulating AI is a challenging task;" *It is difficult for policymakers to assess what AI systems will be able to do in the near future. There is no common framework to determine which kinds of AI systems are even desirable*" (Bhatnagar et al., 2018). While AI holds a significant promise for benefiting humanity, it is crucial to ensure its development is done in a responsible and ethical way.

1.2. Research Purpose and Research Question

The primary aim of this thesis is to investigate the nuances of trust in AI among adults in Sweden, a demographic that is increasingly interacting with AI technologies.[2] Based on the empirical data collected, the study intends to provide a deeper understanding of how various factors such as familiarity with AI and perceptions of task characteristics (objectivity/subjectivity) will influence trust in AI among Swedish adults. The research is centered on the following key questions:

What level of trust do Swedish adults have in AI?

What factors can explain that level of trust?

How do Swedish adults perceive recommendations from AI and humans in different decision scenarios with varying degrees of subjectivity? How can one explain these perceptions?

[2] Due to time constraints the sample in the study does not represent Swedish adults, rather Swedish young adults mainly in the ages 17-35, however throughout the thesis we will refer to the sample as "Swedish adults". This non-representativeness is discussed more in the Methodology.

1.3. Expected Contribution

The primary goal is to deepen understanding of trust in AI, focusing on the factors influencing user trust. Building upon theories and findings in the field of AI and trust, this research aims to contribute to the existing body of knowledge. Our findings delve into the complexities of how individuals perceive and interact with AI, offering insights that are limited not only to academic discussions but also to the practical application of AI in various settings. As Glikson & Woolley (2020) highlighted, AI is increasingly integrating into the daily lives of individuals, making the level of trust put in AI pivotal for shaping organizations‘ futures. This topic is particularly relevant in marketing, as firms need to strategize AI implementation in their operations, considering both customers and employees.

Our study therefore aims to shed light on the dynamics of user trust, providing a foundation for firms to develop marketing strategies that effectively address user concerns and highlight the benefits of AI. This is crucial in an era where AI is becoming increasingly integrated into various products and services. For instance, a global study found that 60% of people are either ambivalent or unwilling to trust AI, highlighting the current state of public trust in AI technologies and underscoring the need for effective trust-building strategies in AI marketing (KPMG, 2023).

Finally, the research‘s implications reach beyond marketing. Policymakers, educators and AI developers can utilize these insights to cultivate trust between users and AI technologies. This understanding is crucial for AI‘s ethical integration and broader societal acceptance.

1.4. Delimitations

Our study‘s data collection was mainly limited to Swedish university campuses, therefore primarily involving students. This decision was made considering the accessibility of respondents and time constraints for the thesis. The collected data can therefore be considered a convenience sample which further limits the research (more discussed in the Methodology). Most respondents were aged 17-35 (91%), with a few older up to 75 years. We retained these older responses, as we deemed them to not significantly impact the study‘s outcomes. Methodologically, we chose a quantitative survey-based experiment which aligns with similar studies conducted. However, it limits our ability to gather detailed qualitative insights into individual AI perceptions and experiences.

Our study focuses on specific AI applications and scenarios we believe many can relate to as well as tasks AI‘s currently can perform or soon will be able to perform; the use of dating apps, academic grading, health diagnoses and career advice (Herrman, 2023). Respondents are asked to rate their level of trust in AI’s performing these tasks, therefore limiting the scope of AI applications that could have been explored, meaning that our findings may not be entirely applicable to other AI contexts.

While our study provides valuable insights into trust in AI among Swedish adults, these findings must be interpreted with an understanding of the delimited geographical focus, the predominantly young demographic and the specific AI scenarios presented.

2. Literature review

*This thesis aims to examine how trust in AI is developed and we have therefore conducted research on previous empirical studies in the area from which hypotheses are formulated. The research was mainly done through SSE Library and Scopus. The following keywords were used; *Trust in AI, *Trust in technology, *Cognitive trust, *Emotional trust, *Familiarity with AI, *Humans and AI, *Technological anxiety and *Task characteristics.*

2.1. Trust in AI

As mentioned in the introduction, AI technologies have developed rapidly in recent years and thus research that is related to the current AI landscape is limited. Nonetheless, insights from previous studies on technological advancements and trust dynamics in AI, even if dated, remain relevant for the sake of this thesis as trust plays a crucial role in the acceptance and adoption of AI technologies, influencing how individuals and organizations interact with and rely on AI systems (Glikson & Woolley, 2020; Hengstler et al., 2016). Firstly, looking at the definition of trust, one of the most cited is by Mayer et al. (1995) who define it as follows; "*The definition of trust proposed in this research is the willingness of a party to be vulnerable to the actions of another party based on the expectation that the other will perform a particular action important to the trustor, irrespective of the ability to monitor or control that other party*". Whether one allows themselves to be vulnerable to the actions of an AI depends on many factors as will be discussed in this literature review. To begin with, trust is often divided into cognitive and emotional trust (Glikson & Woolley, 2020; Johnson & Grayson, 2005; McAllister, 1995).

2.1.1. Cognitive trust in AI

Cognitive trust is rooted in rationality by evaluating the capabilities of the party to be trusted which is affected by factors such as task characteristics (discussed in section 2.4) and reliability (Schoorman et al., 2007). Conclusively, by evaluating AI's functional abilities, competence and performance consistency, cognitive trust is developed. Empirical research on the subject concludes that cognitive trust in AI can be boiled down into the following components; trust trajectory, tangibility, transparency, reliability, task characteristics and immediacy behaviors (Glikson & Woolley, 2020). Moreover, Glikson and Woolley (2020) notes that; *"When researchers examine cognitive trust in AI, they measure it as a function of whether users are willing to take factual information or advice and act on it, as well as whether they see the technology as helpful, competent, or useful"*.

2.1.2. Emotional trust in AI

Emotional trust (also called affective trust), stems from more effective and sometimes irrational factors, including emotions and moods such as gut-feeling (Komiak & Benbasat, 2006; Mcallister, 1995). It relates to users forming a personal bond or sense of security with the technology, often overshadowing aspects like reliability and transparency. Glikson &

Woolley (2020) identifies key factors influencing emotional trust as tangibility, anthropomorphism and immediacy behaviors.

A common misconception is that AI lacks human-like abilities like emotion and creativity, additionally studies indicate higher trust in AI for technical tasks than those requiring social intelligence (Dietvorst, 2016; Gaudiello et al., 2016.; Glikson & Woolley, 2020). However, advancements in generative AI tools now demonstrate AI's capability for creative tasks like music creation, expressing emotions and creating art and architecture, as seen with tools like DALL-E (Ploennigs & Berger, 2023). Distinguishing whether these creations are crafted by an AI or a human is difficult, and such AI tools might therefore pass the Turing test in some cases (Turing, 2009). This makes it intriguing to investigate how cognitive and emotional trust in AI varies with its evolving human-like abilities.

Cognitive trust works as the base for emotional trust, meaning that emotional trust cannot be built without first establishing cognitive trust (Johnson & Grayson, 2005). However, for users to fully embrace AI, both cognitive and emotional trust dimensions are essential. In other words, building real trust involves fulfilling the cognitive aspect, where users trust the technology's reliability, as well as the emotional aspect with the establishment of a personal connection with the user (Dietvorst et al., 2018). As AI becomes more integrated into daily life, the trust level in these technologies is continuously evolving. Positive, personalized AI interactions over time can enhance emotional trust. This highlights the need to design AI systems that are not just reliable and efficient but also foster positive user experiences, nurturing emotional trust.

The literature indicates that while both cognitive and emotional trust are crucial in AI, cognitive trust is likely to be more predominant, especially in the initial stages of AI adoption and interaction. The development of emotional trust in AI, though significant, might follow a more gradual trajectory, influenced by personal experiences and emotional connections with the AI technology. This leads us to the formulation of our first hypothesis:

H1: Individuals are likely to have more cognitive than emotional trust in AI.

2.1.3. The effect of familiarity and age on trust in AI

The relationship between age and trust in AI is key to understanding AI technology adoption. Research indicates age significantly influences trust in AI, with older adults typically having less trust compared to younger individuals (Antes et al., 2021; Gillath et al., 2021). The correlation between an individual's familiarity with AI and their trust level has also been extensively studied. Findings suggest a positive correlation, meaning increased AI exposure tends to build greater trust. This reveals the significance of user experience in AI acceptance and integration. As users become more familiar with AI and its capabilities, their trust in these systems seems to grow, highlighting experience's vital role in shaping AI perceptions (Gillath et al., 2021; Oksanen et al., 2020).

Moreover, a recent study by "Internet Stiftelsen" shows 30% of Swedes used AI applications in the past year, with 60% usage among 18-34-year-olds but only 5% in the 65-84 age group (Svenska Internetstiftelsen, 2023).[3] This further indicates that age may have a negative impact on trust in AI, possibly due to less familiarity with the technology.[4]

2.1.4. Trust trajectory in AI

Research on familiarity's impact on AI trust shows it varies by AI representation, namely embedded, robotic or virtual. In robotics, trust typically starts low and grows with interaction, mirroring human trust development (Glikson & Woolley, 2020). Virtual and embedded AI shows more contradicting findings where trust often starts with being high and then decreases after errors for both representations (Glikson & Woolley, 2020; Hoff & Bashir, 2015).This drop might stem from human-like features creating unrealistic AI expectations, leading to disappointment when unmet (Ben Mimoun et al., 2012; Glikson & Woolley, 2020). This highlights AI's reliability and error handling as crucial in meeting user expectations and forming trust.

On the other hand, for embedded AI, research found that when people were informed about interacting with AI's unfamiliar to them, initial trust tended to be low (Eslami et al., 2015; Möhlmann & Zalmanson, 2017). Similarly, in virtual AI contexts, positive experiences were seen to significantly boost trust, suggesting an initially low trust level (Wang et al., 2016). Additionally, a field study with a virtual museum guide also reflected this trend, showing initial negative sentiment and supporting that initial trust is low (Kopp et al., 2005).

Moreover, research indicates that tangible, human-like features in AI can help establish initial trust, which then grows with user interaction (Looije et al., 2009). In conclusion, AI's consistent and effective error management can notably enhance user trust over time. Recognizing how trust evolves is therefore vital for AI development, emphasizing the importance of creating AI systems that are approachable, reliable and adaptable to user experiences.

It is crucial to note that the findings about different AI representations are from before ChatGPT's release. Since then, trust dynamics in AI are likely to have shifted due to increased user interaction, with limited research available on these changes. We believe that as AI evolves and people become more accustomed to incorporating AI systems into their daily lives, their trust in AI will likely increase.

[3] *Worth noting is that the study done by Svenska Internet Stiftelsen refers to conscious use of AI applications such as ChatGPT or DALL-E, i.e. not embedded AI (*Glikson & Woolley, 2020*).*

[4] *Due to accessibility of respondents and time constraints for the thesis, our sample consists of predominantly individuals in the ages 17-35 (91%). We therefore acknowledge age as an important factor for trust in AI however we will not test it further in this thesis.*

Building on the insights gathered, we arrive at the following, forward-looking hypothesis regarding the relationship between user experience and trust in AI:

H2: There is a positive correlation between experience in using AI and trust in AI.

2.2. Technological anxiety

The interaction between technological anxiety (hereafter named TA) and AI adoption/user behavior is key. Studies on self-service technologies (hereafter named SST) reveal a negative correlation between high TA and technology, meaning increased TA leads to decreased SST use adoption (Gelbrich & Sattler, 2014; Liu, 2012; Meuter et al., 2003). Moreover, lower levels of TA are linked to higher satisfaction, increased reuse and recommendation likelihood, while higher TA levels result in less satisfaction and reduced positive word-of-mouth effect, particularly for initially satisfied SST users (Meuter et al., 2003). Additionally, research shows that forcing SST usage can diminish technology trust (Liu, 2012). Furthermore, Meuter et al. (2003) found TA to be a more accurate predictor of SST usage than demographics like age and gender. The TA-scale measures an individual's level of anxiety or discomfort towards using new technologies and is measured on a 7-point Likert scale. It was originally developed as a computer anxiety scale focusing on personal computers and has since been modified to reflect more general anxiety with all forms of technology (Raub, 1981). Similarly, we will slightly modify the scale for AI anxiety, the rationale for this will be further explained at the end of the section.

2.2.1. AI anxiety

Li & Huang (2020) and Wang & Wang (2022) both explore the multifaceted nature of AI anxiety, but from different perspectives and with varying methodologies. Li & Huang (2020) delves into the underlying factors of AI anxiety, identifying eight primary contributors, for instance privacy violation anxiety, bias behavior anxiety and existential risk anxiety.

Wang and Wang (2022) focus on the development and validation of an AI Anxiety Scale (AIAS), a tool designed to measure AI anxiety quantitatively. Their study introduces a 21-item scale with four factors: learning, job replacement, sociotechnical blindness and AI configuration. This scale was validated through testing for reliability and various forms of validity, with an aim to develop a standardized method for assessing AI anxiety among individuals. However, for reasons that will be explained further down in this section, we deemed the TA scale to be more appropriate.

Asan et al's. (2020) findings highlight how AI's unpredictability in critical areas, such as healthcare, can significantly impact trust. They advocate for optimal trust in AI, warning against both uncritical acceptance and excessive skepticism. They stress the need for fairness and transparency in AI systems to mitigate anxiety and foster balanced trust, showing the importance of addressing anxiety-related factors to improve trust in AI systems. The study

also notes that concerns about AI's performance, such as biases or inaccuracies due to inadequate or subjective data, can worsen this anxiety, negatively impacting the clinician's acceptance and trust in AI technologies, similar to conclusions by Glikson & Woolley (2020).

In measuring anxiety related to AI technologies, we chose between new AI-specific anxiety scales and the established TA-scale. Recent studies like Li et al. (2020) and Wang & Wang (2022) offer insights and tools for AI anxiety but have limitations in that it is still not very established and the two papers are taking very different approaches. The TA-scale, by Meuter et al. (2003) provides a robust, validated framework for predicting technology adoption and user behavior. Its general applicability and effectiveness make it suitable for this study's broader AI focus as questions can be adjusted to the needs of the study - framing questions so they relate both to AI and technology in general. We believe that this choice best supports our research goals and accommodates AI's unpredictable nature, as seen in studies like Asan et al. (2020), justifying the use of the TA-scale to explore the link between TA and trust in AI technologies. This leads us to the following hypothesis:

H3: Higher levels of TA negatively correlate with trust in AI.

2.3. Task characteristics

AI's are believed to be more efficient in some tasks than others. As previously mentioned, more trust is generally put in AI's for technical tasks than those that require social intelligence. Thereby, task characteristics play a crucial role in developing trust in technologies (Hancock et al., 2011). Task characteristics can be divided into two dimensions, subjectivity and objectivity. We follow (Castelo et al., (2019) suggestion on how to define these; "*We define an objective task as one that involves facts that are quantifiable and measurable, compared with subjective tasks, which we define as being open to interpretation and based on personal opinion or intuition*". This is further confirmed in Glikson & Woolley's (2020) empirical review of trust in AI, as studies show how crucial the task is for the emergence of cognitive trust in AI's.

As consumer access to AI applications grows, they increasingly face choices between relying on AI or humans. Even though AI's can outperform humans on several tasks, people often prefer humans, especially in tasks perceived as subjective or intuition-based (Castelo et. al, 2019). This preference relates to emotional and cognitive trust: subjective tasks typically being associated with emotional abilities and objective tasks with cognitive abilities (Inbar et al., 2010). Hence, individuals tend to trust AI more for objective tasks, believing AI is less capable in subjective tasks that require emotional skills. Ultimately, whether a technology eventually is adopted depends on individuals beliefs about how effective it will be (Davis et al., 1989). Thus, we propose the following hypothesis:

H4: Tendency to rely on AI is positively correlated with the perceived objectiveness of the task.

2.3.1. The importance of context

AI's usefulness, fairness and risk vary greatly across contexts, affecting trust. Research indicates that task characteristics are malleable and increasing a task's perceived objectivity through reframing can enhance adoption since the user's confidence in AI's capabilities to handle them effectively increases. Additionally, presenting examples of algorithms with human-like capabilities, like art creation, makes the algorithm appear more competent for such tasks. Lastly, certain tasks are seen as more consequential, meaning that the consequences of performing the task poorly are more serious for some tasks than others, therefore affecting the trust put in AI (Castelo et al., 2019).

Studies also suggest that high self-confidence and belief in one's abilities can hinder technology adoption, as these individuals view themselves as superior to the technology and thus rely less on it (Lewandowsky et al., 2000).This can be compared to experts relying more on their own judgment and being less open to advice from others versus to non-experts (Glikson & Woolley, 2020).

A study by Araujo et al. (2020) shows that trust in AI differs in various areas. AI in healthcare, for instance, is viewed as more useful and fairer than in other sectors, suggesting higher trust where AI's contributions are seen as more vital, highlighting the need to consider the specific context of AI application in its development and deployment.

2.4 Overview of Hypotheses

H1: Individuals are likely to have more cognitive than emotional trust in AI.
H2: There is a positive correlation between experience in using AI and trust in AI.
H3: Higher levels of TA negatively correlate with trust in AI.
H4: Tendency to rely on AI is positively correlated with the perceived objectiveness of the task.

---

3. Methodology

As previously discussed, the aim of the thesis is to investigate the nuances of trust in AI among adults in Sweden. In this section of the thesis, the chosen scientific approach is presented, namely a quantitative approach by doing an experimental survey.

3.1. Scientific Approach to the Research Design

In our research, we undertook an objectivist ontological position in our aim of understanding reality (Bell, 2022), therefore assuming that the phenomena we study exist objectively and that social phenomena exist whether people are aware of them or not, they exist independently of the people who observe them. Consequently, we embraced a positivist epistemological approach, meaning that reality exists objectively and that the appropriate way to gather data is to observe phenomena directly, e.g. using surveys. The logic is deductive, where we aim to frame hypotheses, collecting data to test them and seeking to satisfy or falsify them to understand true statements about reality. This contrasts with an interpretivism approach, which seeks to understand rather than explain human behavior (Bell, 2022).

We utilized a quantitative survey to analyze individuals‘ attitudes towards AI by randomly assigning them to groups with different scenarios so that a comparative analysis could be made. Our empirical research design was informed by our Literature Review, which revealed that most prior studies on trust in AI also used quantitative methods, validating our approach (Gillath et al. 2021).

3.1.1 Alternative Approaches

Alternative methods would be to do a secondary analysis of existing data collected by other researchers or official statistics (e.g. SCB), however this poses several limitations. Firstly, unfamiliarity with the data structure and variable coding could be confusing and time-consuming (Bell, 2022). Secondly, the absence of key variables would be prominent as we wanted to conduct an experiment and be able to compare groups depending on if they got AI/human versions but also if they got a subjective/objective scenario. Thirdly, using such data gives us no control over the data quality. Additionally, rapid AI advancements, meant older data might not accurately reflect current attitudes towards AI, which have evolved with increased awareness and everyday use (as discussed in the Introduction).

Alternatively, we could have conducted qualitative interviews, aligning with a constructionist ontology, i.e. viewing social objects as socially constructed, and an interpretivist epistemology position, where reality is viewed as constituted by human action rather than existing objectively (Bell, 2022; Scotland, 2012). A qualitative research would have provided deep insights into varying trust levels in AI, however it faces the following limitations; the interviewee is likely to be influenced by characteristics of the researcher such as age, gender etc., results are influenced by the researcher‘s interpretation and its challenging to replicate due to the lack of standard procedures (Bell, 2022). We chose a quantitative approach for its standardized nature, enabling clear comparisons between scenarios and groups.

3.2. Pilot study

We conducted a pilot survey using Qualtrics, presenting various scenarios to choose which ones to include in the final survey. This ensured clarity in question formulation and distinctiveness in scenario subjectivity/objectivity while maintaining realistic scenarios. The pilot-survey, involving 19 people, ran from October 18-23, 2023. Due to time constraints, it was not expanded to a larger sample. The scenarios who were deemed most subjective and objective were chosen for the final survey, namely *Dating*, *Math*, *Career* and *Doctor* (see Appendix 1). Feedback from this pilot helped refine scenarios and questions for the final survey.

3.3. Main questionnaire and variables

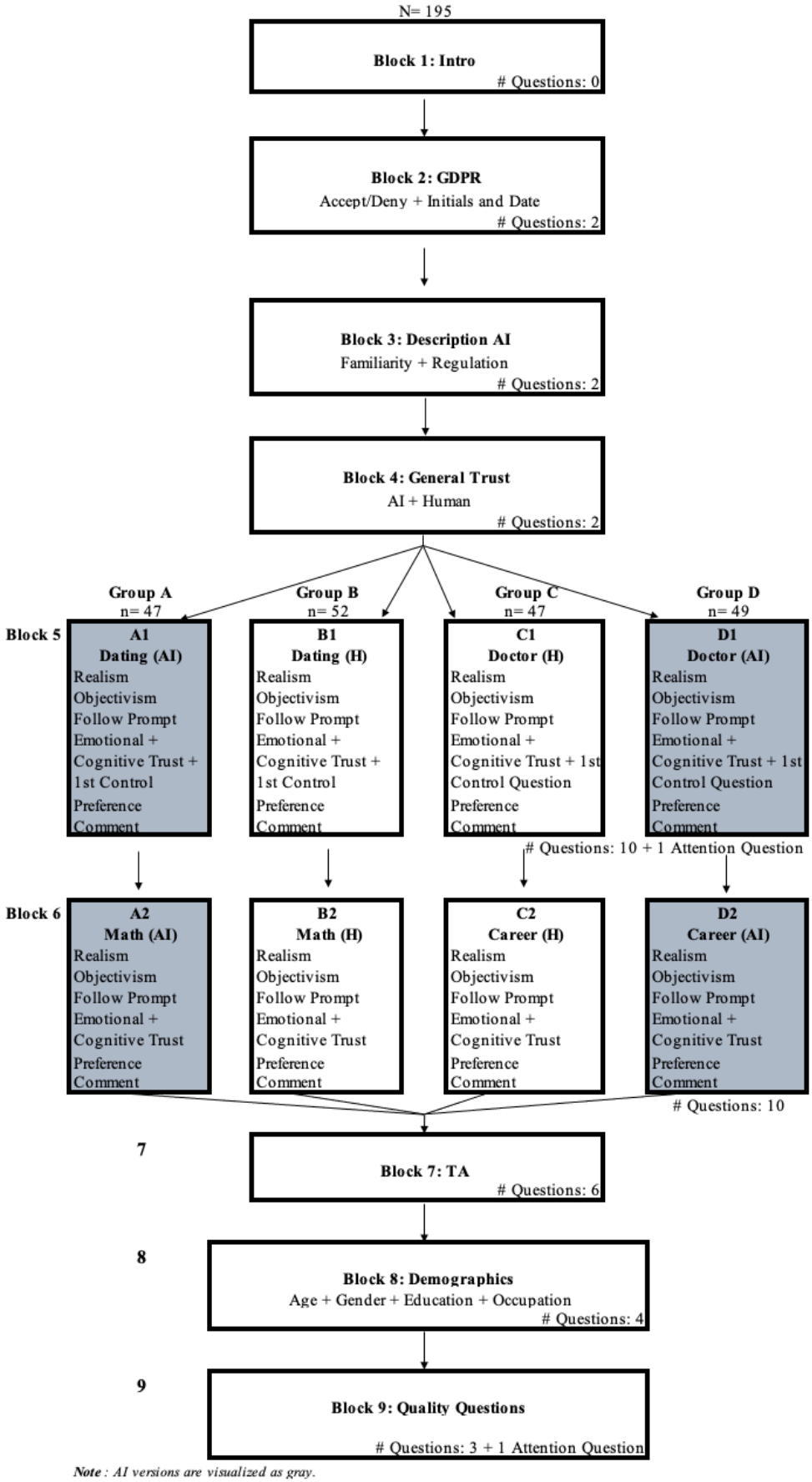
3.3.1. Questionnaire

To minimize misunderstandings, we conducted our survey in Swedish as the target was Swedish adults. The survey comprised 39 questions (excluding comments) and two attention check questions to ensure respondent attention and quality responses (Oppenheimer et al., 2009).[5] In Block 1, an introduction explained the survey's purpose, estimated completion time, contact info and information about how their participation contributed to "Barncancerfonden" (the Children's Cancer Fund) of a 2 SEK donation for every complete response received.[6] In Block 2, GDPR information was presented, requiring initials and date from participants, with those declining redirected to the survey's end. As experience of AI might vary between participants, a definition of AI was presented in Block 3 followed by questions on AI application usage and opinion on AI regulation. Block 4 asked participants to rate their general trust in AI and humans.

The survey featured 4 scenarios, with each participant reading 2. In Block 5, participants were randomly assigned to one of four groups (A, B, C or D) and introduced to the first scenario. Each group received one subjective scenario (*Dating* or *Career*) and one objective scenario (*Math* or *Doctor*), experiencing either *AI* or *Human* versions. Groups A and B received *Dating* and *Math*, while C and D had *Doctor* and *Career*. A and D received *AI* versions and B and C *Human* versions, totaling eight distinct scenarios.[7] Block 6 introduced the second scenario. For both scenarios, participants rated *realism*, *task characteristics*, willingness to *follow recommendation*, *emotional trust*, *cognitive trust* and finally *preference* for AI or human execution. Following scenario 1, respondents were also asked to answer a attention check question. Block 7 focused on *TA*, Block 8 on *demographics* (*age, gender, education* and *occupation*) and Block 9 on survey evaluation and the final attention check question.

[5] *See Appendix 25 for content of the main survey.*
[6] *See section 3.3.2. Survey flow for an overview of the Blocks.*
[7] *To clarify, there were 4 main scenarios (Dating, Math, Doctor and Career), however these scenarios had two versions each (AI or human), therefore this results in 8 distinct scenarios, i.e. A1, A2, B1, B2, C1, C2, D1 and D2 as shown in the survey flow.*

3.3.2. Survey flow



***Note** : AI versions are visualized as gray.*

3.3.3. Variables

This section describes each variable used in our study and how they are measured.[8] In developing the questionnaire, items and measurements were constructed to align with those used in previous studies where possible.

Experience with AI (Independent variable)
This variable measured individuals' usage of various AI applications; large language models, generative AI-images, generative AI-music, voice assistants, voice-language translations, self-driving vehicles, social media, fitness trackers and smart-home devices, chosen for their relevance to current trends. Participants indicated their use of these services, with scores ranging from 0 (no usage) to the maximum of 9 (using all categories), reflecting their experience with AI technologies.[9]

Opinion on AI regulation (Independent variable)
This variable captured participants‘ initial opinions on AI regulation. Respondents were asked to rate their stance on a 5-point Likert scale ranging from “Absolutely not correct” (1) to “Absolutely correct” (5) to the following statement “*I think AI should be regulated*”.

Initial trust in AI (Independent variable)
This variable assessed participants' general trust in AI, taking inspiration from Glikson & Woolley (2020) study on trust trajectory as it allows for a comparison of trust levels before and after exposure to different AI scenarios. It was measured with the question “What is your general level of trust in tasks performed by AI?” answered on a 5-point Likert scale from “Very Low” (1) to “Very High” (5).

Initial trust in Humans (Independent variable)
Similarly, this variable captures respondents' general trust in humans by asking “What is your general level of trust in tasks performed by humans?” using the same 5-point Likert scale as “Initial trust in AI”.

Realism (Dependent variable)
This variable measures the level of realism for scenarios, inspiration was drawn from Dhami et al (2004). Realism in experiments is essential to ensure participant responses mirror real-world behaviors, therefore enhancing experiment validity. If participants see the scenarios as realistic, they are likely to put in more effort and give genuine answers (Dhami et al., 2004). This variable therefore purely worked to ensure validity and is not in focus for the scope of this study. It was measured by asking “How realistic do you deem this scenario to be?” answered on a 7-point Likert scale from “Highly unrealistic” (1) to “Highly realistic” (7).

[8] To ensure the test‘s reliability and validity, all scale measures and questions were presented in Swedish in the survey. This was done considering our target audience of native Swedish speakers, aiming to prevent any misunderstandings and ensure accurate comprehension of the survey questions.
[9] The variable is named “experience” instead of “familiarity” as we deemed experience to better match our way of measuring this variable.

Task characteristics (Dependent variable)
Inspired by Castelo et al. (2019), this variable measured the participants' perceptions of the scenario's level of objectivity/subjectivity. The question was as follows "How objective/subjective do you deem this scenario to be?", answered on a 7-point Likert scale from "Very objective" (1) to "Very subjective" (7).[10]

Follow recommendation (Dependent variable)
This variable was measured using the question "How likely are you to follow the Human/AI's recommendation?" (formulation depends on type of scenario), using a 7-point Likert scale from "Very unlikely" (1) to "Very likely" (7). It assesses participants' tendency to trust and act upon recommendations by AI/humans in different scenarios.

Emotional and Cognitive trust in AI (Dependent variables)
These variables were measured using three questions for each variable, as suggested by Castelo et al. (2019), further confirmed by Johnson and Grayson (2005). It measures the level of emotional and cognitive trust in AI for the specific scenarios on a 5-point Likert scale ranging from "Don't agree at all" (1) to "Completely agree" (5), (see questions in Appendix 29).[11]

AI/Human preference (Dependent variable)
These categorical variable measures whether the respondent prefers AI, humans or has no preference in each given scenario. Participants were asked; "Would you prefer an AI or a very qualified human to perform this task?".

Technological anxiety (TA) (Independent variable)
This variable was measured using 6 out of the 18 questions suggested by Meuter et al., 2003 who we deemed to be most applicable to our study. We modified four questions to be AI-tailored while also keeping two as they are.[12] This was measured on a 7-point Likert scale ranging from "Don't agree at all" (1) to "Completely agree" (7).

Demographics (Independent variables)
The demographic variables assessed included age, gender, educational level and occupation. These measures aimed to deepen our understanding of how demographics correlate with perceptions and beliefs about trust in AI.

[10] Explanations for objectivity/subjectivity and realism were given in the survey to minimize any potential misinterpretations.
[11] Important to note is that due to how the questions were phrased, directions for emotional and cognitive trust differed, to rectify this, emotional trust was reverse coded so that the variables could be compared with 5 indicating high levels of emotional/cognitive trust.
[12] Also here, directions for the questions differed, thus question 4 and 6 (related to TA) were recoded to match the directions of the other questions. We acknowledge that this might have confused respondents.

3.4. Data collection and statistical methods

3.4.1. Data collection

Figure 1. Distribution method

	N = 195	Percentage
Link	n = 50	26%
QR code	n = 145	74%

The survey was conducted via Qualtrics and distributed both physically and digitally from November 6th to 11th, 2023. Physically, it was shared using a QR code at SU, KTH, Stockholm Central Station and Odenplan metro station (74% of responses). Digitally, it was distributed via Facebook and Instagram links (26% of responses). Participants responding via QR code received candy and all were informed about a 2 SEK donation to The Children‘s Cancer Fund for each completed response.

As will be detailed in Results, 84% of respondents were students. Due to time-constraints, this was the most accessible sample and therefore it can be seen as a convenience sample (Bell, 2022). Hooghe et al. (2010) criticize using undergraduate students since they are likely to think and act differently from the general population, e.g. by exerting extra cognitive effort to answer questions “correctly” and higher likelihood being from higher socioeconomic groups. While this limits the representability of our findings, it provides a basis for future research.

3.4.2. GDPR

The General Data Protection Regulation‘s (EUR-Lex, 2016) guidelines were followed in the collection and handling of the empirical data. All data used comes from participants who gave their consent to participate in the study and no data was collected for those who didn’t give consent. We solely gathered information that was required for the sake of the study such as age, gender, education, occupation and initials. Moreover, GDPR-regulations were presented to all participants and all data will be deleted upon completion of the revised thesis beginning in 2024.

3.4.3. Data quality

461 individuals entered the survey and 257 (56%) completed all questions. After removing incomplete responses in Excel, i.e. respondents who failed attention checks and/or didn’t comply with GDPR, 195 responses remained. This means 76% of the complete responses adhered to GDPR, provided initials/date and passed both attention checks. 60 participants failed in one or more of these aspects. Distribution-wise, 30% of responses were collected from either social media or at Odenplan/the Central station, 67% from SU and 4% from KTH

(see Appendix 2 and 3). The 195 responses were allocated as follows: 47 to group A, 52 to group B, 47 to group C and 49 to group D.[13]

3.4.4. Data analysis

We analyzed the data using Excel and R. We derived descriptive statistics (means, standard deviations), conducted a one-way ANOVA, individual t-tests, correlation matrices as well as regressions to analyze our findings. Cronbach‘s alpha was calculated for multi-item questions to evaluate internal consistency and variance inflation factor to test for multicollinearity (Bell, 2022).

3.5. Reliability and validity

Reliability and validity are related to each other as validity presumes reliability, in other words, if a measure is not reliable it cannot be valid (Bell, 2022). Reliability refers to the consistency of measurements, assessing if repeated measurements of the same thing yield consistent results while validity refers to how well a measurement reflects the phenomenon it is intended to measure. Important to note is that high reliability does not necessarily mean that validity is high (Söderlund seminar, 2023).

3.5.1. Reliability

To assess internal reliability, we calculated Cronbach‘s alpha for Emotional trust, Cognitive trust, TA and quality questions. Internal reliability indicates whether a measure consistently evaluates the same concept; lack of it suggests the measure might assess multiple aspects, impacting validity. A Cronbach‘s alpha above 0.7 is normally considered acceptable (Joseph et al., 2019; Bell, 2022). Low alpha values might stem from few questions or weak item correlations (Tavakol & Dennick, 2011). Due to the study‘s anonymous one-time nature, we could not measure stability via test-retest or apply inter-rater reliability. Chronbach’s alpha for all multi-items were above (or close to) 0.7, i.e. there is high internal reliability (see Appendix 4).

3.5.2. Validity

Validity has to do with whether a measure of a concept really measures that concept (Bell, 2022). Face validity was established through discussions with our mentor Patric Andersson, who has extensive experience in conducting research, who could ensure that our measures seemed to reflect the concept concerned. We ensured content validity by utilizing established scales and questions from previous research regarding AI, technology and trust (as discussed in the Literature review and in section 3.3.3).

Looking at construct validity, our measures relate to variables and theories compatible with previous studies such as the negative effect of TA on trust in technologies, higher levels of objectivity for a task having a positive effect on trust in AI as well as findings stating that

[13] *The initial intention was to compare respondents' depending on where they were collected, however as a large majority comes from SU (67%), we decided not to do this analysis.*

individuals have more cognitive than emotional trust in AI (as presented in the Literature review and further discussed in the Results).

The external validity of our findings is limited due to our sample‘s demographic bias towards students, females and adults, necessitating further research with more diverse groups. However, it is strengthened by the high levels of realism related to scenarios reported by respondents (see figure 7). Thus, while our study demonstrates an overall strong validity, future research should put emphasis on external validity to allow for a better representation of how a larger demographic views trust AI (Bell, 2022).

Attention checks

Krosnick (1991) highlighted that surveys often require significant cognitive effort leading to participants choosing the first reasonable option or, in extreme cases, answering randomly, thus reducing an experiment‘s power (Krosnick, 1991; Oppenheimer et al., 2009). To identify such participants, we used attention check questions (also called Instructional Manipulation Checks) embedded within the other survey questions. Moreover, we aimed to mitigate this effect by formulating realistic and interesting scenarios such that respondents would be motivated to put in the effort while answering the survey answer genuinely and with effort. Additionally, the randomization of groups and scenarios prevented friend groups from discussing and influencing each other's answers.

3.5.3. Survey evaluation

The last section (Block 9) of our survey assessed its perceived quality using a 5-point Likert scale. This aimed to gauge participants‘ views on the clarity, relevance and their understanding of the study‘s purpose. Results showed that 89% understood the study‘s purpose, 82% found the questions clear and 93% felt the questions were relevant (see figure 2).

Figure 2. Survey evaluation

	Don't agree	Partly disagree	Neutral	Partly agree	Agree a lot
I understood the purpose of this study.	1%	2%	8%	44%	45%
The questions in the study were clearly formulated	1%	10%	8%	39%	42%
The questions in the study felt relevant to the subject.	0%	0%	7%	32%	62%

***Note**: N=195*

4. Results and analysis

This section of the thesis presents empirical findings of the study, firstly presenting descriptive statistics, secondly an ANOVA analysis, thirdly correlation matrices and finally a regression analysis. We aim to explore if there are any differences in variables depending on if the group got an AI or human scenario and if it was a subjective or objective scenario.

4.1. Descriptive statistics

4.1.1. Sample demographics

Figure 3 gives an overview of the demographics of the sample showing *age*, *gender*, *education* and *occupation* (N = 195). All groups had about the same age distribution with means ranging from 25 to 28 years (SD ranging from 8.52 to 12.52) and a large majority (91%) of respondents being in the ages of 17-35 (see figure 4 for visualization).[14] Furthermore, a majority of respondents (84%) were students, 13% were working part/full-time, 2% pensioners and 1% entrepreneurs. Moreover, the respondents consist of 68% females and 32% men which might cause the result to be skewed. We are aware that this does not represent the Swedish population and the limitation of generalization has been discussed more in the Methodology.

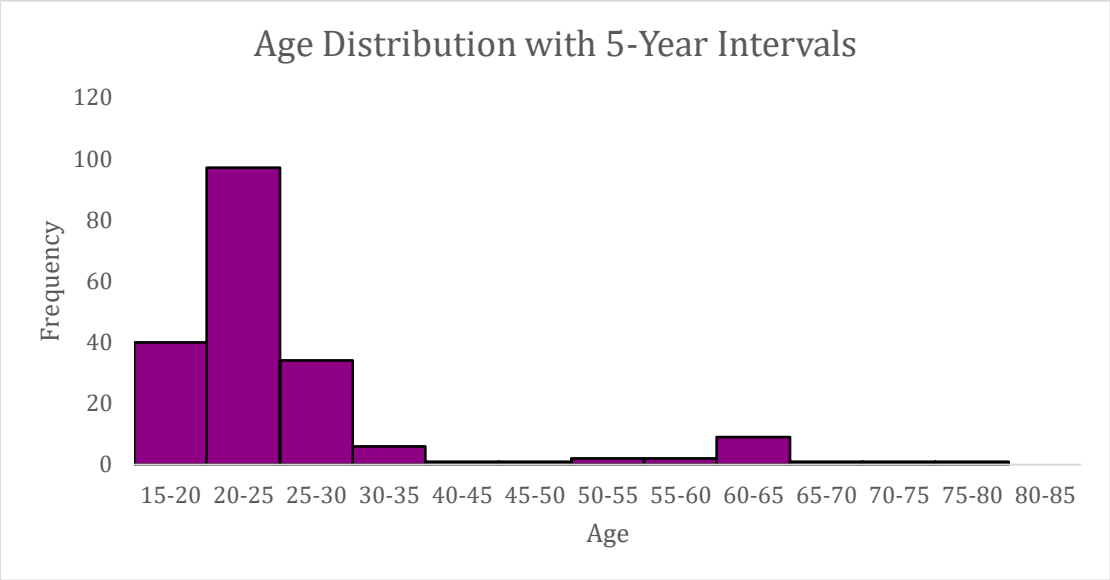
Figure 3. Demographics

	n= 47	*n= 52*	*n=47*	*n= 49*	*N= 195*
Demographics	**A**	**B**	**C**	**D**	**All groups**
Age					
Min	19	17	18	18	17.00
1st Quartile	20	21.75	21	21	20.75
Median	22	23.5	22	23	22.00
Mean	24,6 (8.52)	28,37 (12.52)	26,13 (10.62)	25,43 (8.91)	26.1 (10.33)
3rd Quartile	24	27.25	26	26	26.31
Max	57	68	75	57	75.00
Gender					
Male	32%	44%	28%	22%	32%
Female	68%	56%	72%	78%	68%
Education					
High School	4%	10%	6%	8%	7%
University (not degree)	74%	65%	72%	59%	68%
Bachelor	17%	17%	13%	20%	17%
Master or higher	4%	8%	9%	12%	8%
Occupation					
Student	89%	81%	77%	88%	84%
Part/Full-time	11%	13%	19%	10%	13%
Entrepreneur	0%	2%	2%	0%	1%
Unemployed	0%	0%	0%	0%	0%
Pensioner	0%	4%	2%	2%	2%

***Note**: SD in paranteses*

[14] *As age is highly concentrated, it will not have a significant effect on the results, therefore t-tests were not done to compare age between groups.*

Figure 4. Histogram showcasing the age distribution.



4.1.2. General attitude towards AI

To ensure that there were no systematic differences between groups, questions regarding respondents' *experience of using AI*, their *opinion on AI regulation*, *general trust for humans/AI's* and levels of *TA* were asked. This way we could make sure that all groups had about the same attitude towards AI, not having one group being "anti-AI" for example as that could have an impact on validity. As can be seen in Figure 5, all groups had similar levels of attitude for these variables.[15] We can therefore, as expected, conclude that there are no systematic differences between groups as they had not yet been exposed to experiments. The mean for *experience* scored between 3.53 to 4.13 (SD ranging from 1.53 to 1.76), meaning that respondents reported using on average about 4 out of 9 AI applications provided in the survey (see Figure 6 for overview of which applications were most used). The *opinion on regulation of AI* scored high ranging from 3.94 to 4.28 (SD ranging from 0.79 to 0.91) indicating that the general opinion is that AI indeed should be regulated. Moreover, initial general trust in AI was neutral as means ranged from 2.90 to 3.16 (SD ranging from 0.74 to 0.90) while general trust in humans ranged from 3.58 to 3.66 (SD ranging from 0.60 to 0.78). Lastly, all groups reported about the same levels of TA, ranging from 2.79 to 3.07 (SD ranging from 1.52 to 1.72), all below 4 (neutral) indicating that our sample reports having quite low levels of TA. The question "I prefer using technology that does not involve AI" had the highest level of TA for all groups while questions regarding technological skills, understanding and keeping up with technological advancements had the lowest level of TA, indicating that our sample had rather high confidence in their technology skills.

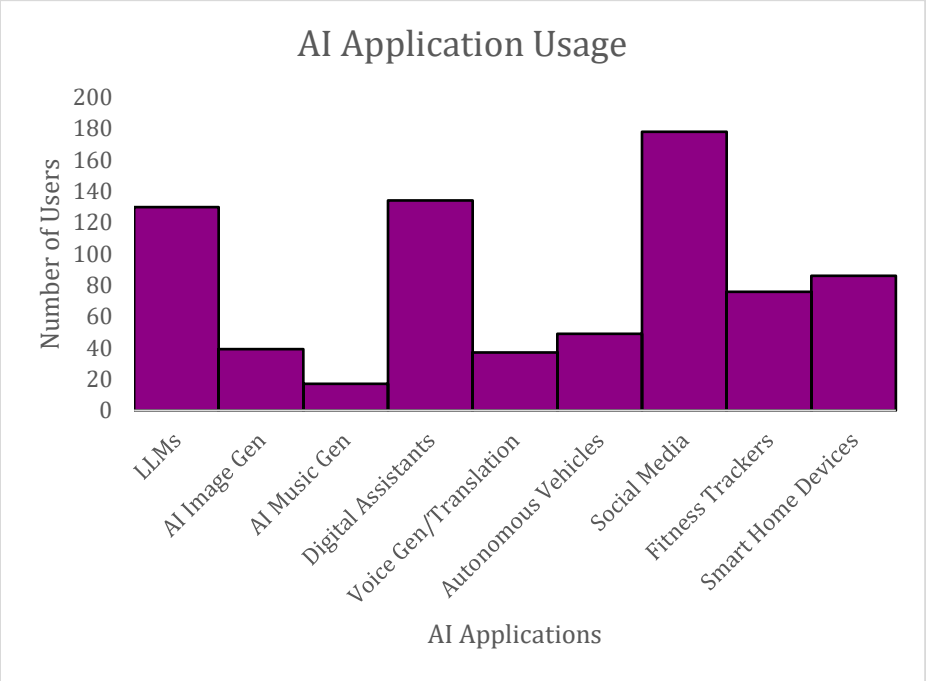
[15] *A one-way ANOVA was conducted for these variables which showed no significant difference for experience, regulation, general trust AI and human between groups (see Appendix 24).*

Figure 5. Levels of experience and initial attitude in AI (see Methodology for explanation of variables)

	n= 47	*n= 52*	*n=47*	*n= 49*
	A	**B**	**C**	**D**
Experience	4.04 (1.76)	3.85 (1.53)	4.13 (1.59)	3.53 (1.72)
Regulation	4.28 (0.79)	4.19 (0.86)	3.94 (0.91)	4.18 (0.81)
General Trust AI	2.98 (0.74)	2.90 (0.85)	3.13 (0.76)	3.16 (0.90)
General Trust Human	3.66 (0.63)	3.58 (0.64)	3.66 (0.78)	3.63 (0.60)

***Note** : SD in paranteses*

Figure 6. Overview of which AI applications participants used



4.1.3. Dependent variables

Moreover, we compared the variables connected to each scenario (dependent variables), i.e. *task characteristics, follow recommendation, emotional trust, cognitive trust and preference,* mainly by looking at means and conducting t-tests to see if they differ significantly shown in Figure 7.[16] To facilitate for the reader, we provide results from one scenario at a time rather than one variable.[17]

Dating scenario *(Group A and B scenario 1)*

Our study reveals that Group A (M = 3.60) and Group B (M = 5.42) significantly differed in their ratings of task characteristics ($p < 0.001$). Despite the task‘s intended subjectivity (as explained in the Methodology), Group A rated it as more objective. Group B (M = 4.17) was

[16] *As discussed in the methodology, realism was not a variable of focus for the scope of this study and therefore not discussed in this section. However, all scenarios scored high with means ranging from 4.49 to 6.19, for more insights see Figure 7.*

[17] *As this study contains an extensive amount of data and comparisons, SD is not presented in the running text, please see Figure 7 for more insights.*

significantly more likely to follow the recommendation than Group A (M = 3.53), i.e. respondents were more likely to follow the human recommendation (p = 0.032). Both groups exhibited similar levels of emotional (Group A: M = 3.11; Group B: M = 3.01) and cognitive trust in AI (Group A: M = 3.52; Group B: M = 3.55), with cognitive trust significantly higher than emotional trust in both groups (Group A: p= 0.037; Group B: p= 0.002), which supports H1.[18] Preference for a human to perform this task was 66% in Group A and 75 % in Group B, while AI preference was 9% in group A and 6 % in group B.

Math scenario (Group A and B scenario 2)
Both Group A (M = 2.45) and Group B (M = 3.25) perceived this scenario as objective, but group A rated it significantly more so (p < 0.001). Also here, Group B was significantly more likely to follow the recommendation (Group A: M = 4.32, Group B: M = 5.35; p = 0.007). Emotional trust in AI was similar for both groups (Group A: M = 3.79, Group B: M = 3.55). However, cognitive trust was higher in Group B (M = 4.17) than in Group A (M = 4.17; p = 0.038), with cognitive trust significantly higher than emotional trust only for Group B (p = 0.025). Both groups had about the same *preference* for humans (Group A: 60%; Group B: 62%).

Doctor scenario (Group C and D scenario 1)
Both groups rated the scenario to be more objective than subjective, however group D rated it as significantly more objective (Group C: M = 3.45 Group D; M = 3.02; p < 0.001). Follow recommendation was rather high for both groups and not significantly different (Group C: M = 4.98; Group D: M = 4.69) meaning that for this task, our sample would trust the diagnosis to be correct about as much if it was a human as if it was an AI making the diagnoses which relates to H4. However, *emotional trust* in AI was significantly higher for group D (Group C: M= 3.03; Group D: M = 3.48; p = 0.020). Cognitive trust in AI was also significantly higher for group D (Group C: M = 3.90; Group D: M= 4.18; p = 0.083). In both groups, cognitive trust significantly exceeded emotional trust (p < 0.001 for both groups). Finally, *preference* for humans was 79% for C and 69% for D, meaning that the human version scenario got higher preference for humans.

Career scenario (Group C and D scenario 2)
Group C rated the scenario as significantly more subjective than group D (Group C: M = 4.57; Group D: 3.80; p = 0.014). Follow recommendation was higher for group C (Group C: M = 4.45; Group D: M = 4.00), however not significantly different. Emotional trust was higher for D than C (Group C: M = 3.58; Group D: M = 3.70), the same for cognitive trust (Group C: M = 3.39; Group D: M = 3.84), however none of them are significantly different. Moreover, for this scenario cognitive and emotional trust did not differ that much in contrast to the other scenarios, for group C emotional trust was even higher than cognitive. Finally, this scenario was the one with the lowest preference for humans of all scenarios, which somewhat contradicts the literature that suggests subjective tasks have a high preference for humans over AI‘s.

[18] *See Appendix 22 for an overview of p-values between emotional and cognitive trust.*

Figure 7. Means and SDs for scenario-variables

	Dating		Math		Doctor		Career		
Means for variables (SD)	A1	B1	A2	B2	C1	D1	C2	D2	
Realism	5.02 (1.57)	5.02 (1.23)	$4.49\ (1.90)_{a}$	$6.19\ (1.14)_{a}$	5.94 (1.29)	5.63 (1.36)	5.47 (1.56)	5.45 (1.60)	
Task characteristics	$3.60\ (1.48)_{b}$	$5.42\ (1.24)_{b}$	$2.45\ (1.38)_{c}$	$3.25\ (1.58)_{c}$	$3.45\ (1.46)_{d}$	$3.02\ (1.39)_{d}$	$4.57\ (1.58)_{e}$	$3.80\ (1.67)_{e}$	
Follow recommendation	$3.53\ (1.43)_{f}$	$4.17\ (1.50)_{f}$	$4.32\ (1.99)_{g}$	$5.35\ (1.75)_{g}$	4.98 (2.08)	4.69 (1.47)	4.45 (1.30)	4.00 (1.27)	
Emotional trust in AI	$3.11\ (1.15)_{p}$	$3{,}01\ (1.02)_{k,\,q}$	3.55 (1.09)	$3.79\ (1.01)_{k,\,r}$	$3.03\ (1.05)_{h,l,\,s}$	$3.48\ (0.85)_{h,\,m,\,t}$	$3.58\ (0.90)_{l}$	$3.70\ (0.90)_{m}$	
Cognitive trust in AI	$3.52\ (0.71)_{p}$	$3.55\ (0.72)_{n,\,q}$	$3.84\ (0.76)_{i}$	$4.17\ (0.65)_{i,\,n,\,r}$	$3.90\ (0.83)_{j,\,s}$	$4.18\ (0.71)_{j,\,o,\,t}$	3.39 (0.83)	$3.84\ (0.71)_{o}$	
Preference									**All groups**
Human	66%, n= 31	75%, n= 39	60%, n= 28	62%, n=32	79%, n=37	69%, n=34	55% n=26	51% n=25	64.62% n = 252
AI	9%, n= 4	6%, n=3	17%, n= 8	19%, n=10	6%, n=3	8%, n=4	15% n=7	12% n=6	11.54% n = 45
No preference	26%, n= 12	19%, n=10	23%, n= 11	19%, n=10	15%, n=7	22%, n=11	30% n=14	37% n=18	23.85% n = 93

***Note** : Letter stands for group, number stands for scenario. Ex. A1 = scenario1 for group A.*
Group A and D got AI scenarios (marked in gray colour).
SD stands in paranteses.
Means that share subscripts differ at p <0.05
Preference all groups: Every respondent answered this question twice in total as the question was presented after each scenario. This means that the total n is double the amount of respondents.

In conclusion, these comparisons suggest that directly introducing a task with a qualified AI, increases the perceived objectivity of the task as all AI scenarios were rated significantly more objective than their human counterparts. This supports the findings from Castelo et al. (2019), as discussed in the literature review. All scenarios (except cognitive trust for the dating scenario) received higher levels of emotional and cognitive trust for the groups who read the AI versions (although not significant in all situations). Moreover, our findings suggest that people are more likely to follow the recommendation of a human than an AI (although not significant for all scenarios) as well as to prefer humans over AI‘s, relating to the research by Glikson & Wolley (2020). Furthermore, overall cognitive trust in AI was higher than emotional trust in AI (although not significant for all scenarios), further aligning with the findings by Castelo et al. (2019). These findings will be further discussed in the coming sections.

4.2. ANOVA analysis and t-tests

To examine our study on a group level, our initial statistical approach included a one-way Analysis of Variance (ANOVA), supplemented by post-hoc Tukey‘s Honestly Significant Difference (HSD) test. This combination of statistical methods was chosen for its robustness in comparing mean scores across the four different groups within our experiment (Bell, 2022). For the ANOVA analysis, tests will be considered significant if their p-values are less than 0.05.

4.2.1. ANOVA results

Our analysis involved a one-way ANOVA to assess group-level differences, focusing on one independent variable (Group A-D). This was complemented by Tukey‘s Honestly Significant

Difference (HSD) test for its effectiveness in comparing mean scores across four groups.[19] The results of the ANOVA provide insights into the impact of our independent variable, the group/scenario combinations, on the dependent variables; *task characteristics*, *follow recommendation*, *emotional trust* and *cognitive trust*.

Figure 8. ANOVA results

Variable	SumSq	Mean Sq	F-value	Pr(>F
Task characteristics	102.80	34.28	12.75	< 0.001 **
Follow recommendation	43.30	14.45	5.17	< 0.005 **
Emotional Trust	4.90	1.65	1.58	0.195
Cognitive Trust	5.36	1.79	2.95	< 0.005 **

As can be seen in Figure 8, the ANOVA analysis indicates significant group differences for all dependent variables apart from emotional trust. Given these significant findings, a post-hoc Tukey test was conducted for all variables, except for emotional trust, in order to determine for which groups the variables differ. We chose to focus on comparing Group A and B together and Group C and D together to allow for just limiting ourselves to one changing variable (if groups received the AI or human version of the scenario). See results for Tukey‘s HSD test in Appendix 23.

4.2.2. Post-hoc results - Tukey‘s HSD test

The perception of *task objectivity* differed significantly between groups in certain contexts. In the Math and Dating scenarios, Group A perceived tasks as more objective compared to Group B ($M = 3.03$ vs $M = 4.35$; $p < 0.001$). However, in professional contexts such as doctor and career advice, no significant difference was observed between Group C and Group D ($M = 4.01$ vs $M = 3.41$; $p = 0.055$).

There was a notable difference for *follow recommendation*. Group B demonstrated a higher inclination to follow human advice as opposed to Group A, which was more inclined towards AI recommendations ($M = 4.76$ vs $M = 3.93$; $p = 0.003$). Contrastingly, in professional settings, no significant difference was detected between Group C and Group D ($M = 4.72$ vs $M = 4.35$; $p = 0.423$), suggesting a comparable level of trust in both AI and human advice.

In terms of *cognitive trust* in AI, there were no significant differences between the groups. Group A and Group B exhibited similar levels of cognitive trust ($M = 3.68$ vs. $M = 3.86$; $p = 0.357$). Likewise, no significant difference was found between Group C and Group D ($M = 3.65$ vs. $M = 4.02$; $p = 0.745$), indicating that the type of agent (AI or human) does not significantly influence trust.

[19] *While initially considering a Bonferroni correction, we opted for Tukey‘s test due to its ability to handle multiple comparisons and control Type I errors effectively (Reed, 2003).*

The study indicates that for personal or social interactions, people prefer human advice over AI. However, when it comes to tasks that require a high level of expertise, such as medical diagnoses or career planning, both AI and humans are viewed similarly.

4.3. Correlation

In order to further understand the relationship between our variables, how they correlate and to find patterns, we conducted one correlation matrix for each scenario (eight in total) containing the following variables: *Task characteristics, Follow recommendation, Emotional trust, Cognitive trust, TA, Gender, Experience using AI, Opinion on regulation, Initial trust in AI* and *Initial trust in humans* (see Appendix 5-12). We created eight matrices since we wanted to take the type of scenario as well as if respondents got human or AI versions into consideration.

The matrices show that experience had a positive correlation with both emotional and cognitive trust for all scenarios, however it was only significant in 4 out of 16 cases (25%). Moreover, experience in using AI correlates positively with initial trust in AI for all groups and is significant in 2 of 4 cases. Furthermore, we can see that experience in using AI correlates negatively with TA for all 4 groups and is significant in 3 out of 4 groups (75%), meaning that those with more experience with AI applications have lower TA. Moreover, TA significantly negatively correlates with emotional and cognitive trust for all groups and scenarios (except for cognitive trust in A2 and D2). Furthermore, follow recommendation positively correlates with cognitive trust in all scenarios (except for B1 where the coefficient is -0,01), and is significant in 4 out of 8 scenarios (50%). Initial trust in AI correlates positively with emotional and cognitive trust in AI for all cases and is significant in 7 of 16 cases.[20]

Moreover, correlation matrices were also conducted in between subjects comparing variables for each group (A-D) between scenario 1 and 2 (see Appendix 13-16). This way, it is possible to explore if individuals‘ answers in the first scenario correlates with the answers in the second scenario provided. We therefore look at variables that are affected by type of scenario; *task characteristics, follow recommendation, emotional trust* and *cognitive trust.*

Firstly looking at task characteristics, we can conclude that overall, answering high levels of subjectivity in the first scenario does not correlate with answering high subjectivity in the second, except for group A. The same goes for *Follow recommendation,* which is solely positively significant for group C, meaning that here individuals think similarly regarding following recommendation in both scenarios. Thirdly, emotional trust is positively significant for all groups, answering higher emotional trust in one scenario seems to correlate with answering higher emotional trust in the second one. Lastly, cognitive trust for both scenarios positively correlates with each other, however only significantly for group C and D, i.e. those who got the “Doctor” and “Career” scenarios. In conclusion, those variables that positively

[20] *Correlations coefficients are not mentioned in the running text as it would be too extensive to report for all matrices, see Appendix 5-12 for more insights.*

correlate with each other means that individuals were having similar opinions in scenario one and two.

4.4. Regression analysis

In order to further test our hypotheses, multi regression analyses were conducted for the variables that we wanted to test for our hypotheses. We did this for each scenario and for three dependent variables; *emotional trust*, *cognitive trust* and *follow recommendation*, in other words 12 multi regressions in total (see Appendix 17-19). The independent variables included in the model were *AI/human, task characteristics, experience using AI* and *TA.*[21] The equation looks as follows.[22]

$$DV = \beta_0 + \beta_1 (AI/Human) + \beta_2 (Task\ characteristics) + \beta_3 (AI\ experience) + \beta_4 (TA) + \mathcal{E}_i$$

The main reason for doing a multi regression analysis is that it tests several variables simultaneously and shows which variables are the most important for predicting the dependent variable (Bell, 2022).

Firstly, looking at the regression on emotional trust, a majority (75%) of the scenarios have a negative Beta for AI/human meaning that getting the human version scenario resulted in less emotional trust in AI. However it was only significant for the Doctor scenario, $r(91) = -0.41$, $p < 0.05$. Furthermore, the Beta for TA was highly significant and negative for all scenarios, meaning that reporting higher levels of TA significantly lowered emotional trust in AI (Beta ranging from -0.32 to -0.52). Objectivism and experience did not have a significant effect on emotional trust in this regression. R2 ranged from 0.20 to 0.34.

Moving on to the regression on cognitive trust, AI/Human had a significant effect on the Math and Doctor scenario such that getting a human in the Math scenario significantly increased the level of cognitive trust in AI while getting a human in the Doctor scenario significantly lowered the cognitive trust in AI. Moreover, we see that objectivism was significant for the math and doctor scenario with negative betas (although very low for the Doctor scenario with $b = -0.089$), meaning that rating these scenarios as more subjective leads to less cognitive trust in AI which relates to the findings by Castelo et al. (2019). TA was also significant for all scenarios with a negative Beta meaning that as TA goes up, cognitive trust goes down. R2 ranged from 0.10 to 0.21.

Lastly looking at the regression on follow recommendation, AI/human was positive and significant for the Math scenario, meaning that reading the human scenario led to significantly higher likelihood of following the recommendation. Moreover, experience with AI was significant with a positive beta for the Doctor scenario, indicating that higher AI experience increases the likelihood of following the doctor or AI‘s recommendation. R2 ranged from 0.04 to 0.11.

[21] *AI/human is a binary variable with AI= 0, Human= 1.*
[22] *DV= Dependent variable, i.e. emotional trust, cognitive trust and follow recommendation.*

Moreover, multicollinearity was tested for all regressions since it can affect regression results and the reliability. In a multiple regression model, multicollinearity occurs when there is a correlation between several independent variables. The variance inflation factor (VIF) for all models ranges between 1.03 to 1.49 (see Appendix 20) which is considered low and positive for the model's reliability, suggesting greater independence among variables (Alin, 2010).

4.5. Implications for Hypotheses

H1: Individuals are likely to have more cognitive than emotional trust for AI.

Upon conducting t-tests testing for significance between emotional/cognitive trust and taking mean across all scenarios, cognitive trust in AI was higher than emotional trust in 7 out of 8 of the cases (significant in 5 out of 8 cases), thus we have empirical support for H1.

Empirical support found

H2: There is a positive correlation between experience in using AI and trust in AI.

Experience in using AI was not significant in predicting cognitive or emotional trust in any of the scenarios in the regressions and only significant in 25% of cases for the correlation matrices. We therefore lack enough empirical support for H2.

Lacks enough empirical support

H3: Higher levels of technological/AI anxiety negatively correlate with trust in AI.

Looking at TA, the regression analysis reveals that TA had a negative Beta and was significant in predicting emotional and cognitive trust for all scenarios. Additionally, the correlation matrices shows that in 14 of 16 cases TA significantly correlates negatively with trust in AI. Thus we have empirical support for H3.

Empirical support found

H4: Tendency to rely on AI is positively correlated with the perceived objectiveness of the task.

When viewing task characteristics and comparing human and AI scenarios in terms of significance (t-test) and mean, all AI scenarios were rated significantly more objective than their human counterparts. Thus, we have empirical support for H4.

Empirical support found

---

5. Discussion

This part of the thesis discusses the results presented in the previous chapter and what conclusions we can draw from the findings of the study. We then present implications, limitations, a general discussion and suggestions for future research.

5.1. Key findings

5.1.1. Descriptive, demographics, familiarity and opinion on regulation

In our study, the majority of respondents were young Swedish adults, with 68% female and 32% male, mostly aged 17-35 years (average age 26). As 84% were students, our findings predominantly reflect adults‘ views in a university context. The participants showed moderate engagement with AI, using an average of 4 out of 9 AI tools, with no significant differences between groups. On AI regulation, there was strong agreement, scoring between 3.94 to 4.28 on a 5-point Likert scale, indicating Swedish adults‘ awareness and concern for oversight and ethical considerations in AI deployment and use.

5.1.2. Research questions

In terms of level of trust, Swedish adults initially reported moderate AI trust, with group means ranging from 2.90 to 3.16 on a 5-point Likert scale. Yet, the study found higher trust levels for tasks in the scenarios, with all averaging 3 or above in both emotional and cognitive trust. This suggests task nature significantly affects AI trust, supporting Castelo et al. (2019). Looking at how Swedish adults view the different recommendations, a consistent human preference over AI is found from higher scores for tasks completed by humans. A significant number of participants indicating “no preference” suggests openness to AI in decision-making, a point we will delve into in our explorative discussion. Respondents exposed to AI scenarios perceived them as more objective and expressed higher AI trust levels, suggesting direct AI involvement in a task enhances its perceived objectivity and trust. This once again supports findings from Castelo et al. (2019), indicating less trust in algorithms for tasks viewed as subjective.

5.1.3. Other implications

The main implication of our study is to contribute to how marketers and organizations can use our findings to draw conclusions on how to increase trust in AI in order to increase adoption of AI. This is especially important as the AI landscape is rapidly evolving and organizations (and individuals) not adopting the technology risk being left behind (Glikson & Woolley, 2020; USPTO, 2020). Our study indicates that one way to increase trust in AI is to focus on presenting AI‘s in a manner that emphasizes the objectivity of the task and qualifications of the AI, similar to how the scenarios were perceived as more objective when introduced with a qualified AI. Moreover, efforts to reduce TA can be done to increase trust in AI as our study finds TA to negatively correlate with trust in a significant way. Moreover, as our study found that cognitive and emotional trust in AI differs significantly, while

previous literature stresses the importance of establishing both dimensions of trust in order to develop long-term trust, there is a clear need for AI systems to improve on factors like empathy, personalization and emotional intelligence, especially as advancements are moving towards cultivating personal relationships with AI (The New York Times, 2023).

Connecting these findings to the strong agreement that AI should be regulated, regulators should aim to develop guidelines that address technical reliability and ethical usage, including data privacy, algorithm transparency and fairness. This builds cognitive trust while also nurturing emotional trust over time. The varied nature of tasks means that regulations should be tailored specifically to each type of task.

5.2. Limitations

Our thesis, like any empirical research, faces limitations. As mentioned in the methodology section, time constraints led us to focus primarily on a young demographic, predominantly students (84%), with a notable gender imbalance (68% female) and a limited sample size (N = 195). This demographic specificity restricts the applicability of our results to a wider, more diverse population, potentially biasing the findings toward the perspectives and AI trust levels of this group.

The fast pace of AI development and its broad scope might have led to differing interpretations among users, as seen in respondents‘ varied comments in our questionnaire. For instance, some believed current AI tools could not correct a math exam, while others saw them as efficient. This variation reflects the ongoing lack of clarity and consensus on defining AI (O’Shaughnessy et al., 2023). We attempted to address this by providing a general AI definition inspired by Gillath et al. (2021). Additionally, we acknowledge the potential for experimenter demand effects (EDE), as discussed by Zizzo (2011), where our scenario design and context might have subtly swayed participants‘ responses to align with perceived study objectives. It is also crucial to note that our study measures only intentions, not actual behavior.

Another limitation concerns respondent attention. Despite instructions to imagine scenarios, some comments like “I do not use dating apps” suggest misunderstandings, potentially impacting the study‘s reliability (Oppenheimer, 2009). Additionally, the rapid rise of AI as a dominant societal topic, influenced by media reports on incidents like self-driving car accidents or Sam Altman‘s departure from OpenAI, might have affected responses, particularly regarding AI regulation and trust (New York Post, 2023; Göteborgs-Posten, 2023). The widespread recent discussion on AI‘s legal and ethical aspects could also have influenced opinions on AI regulation and trust. A more detailed exploration of respondents‘ AI usage experience might have provided a more accurate measure of the experience-trust relationship, a point further elaborated in “Suggestions for future research”.

5.3. Explorative discussion

5.3.1. Segmentation based on gender

Though not the primary focus of our study, interesting gender-related patterns emerged in our data. Our correlation matrices showed that emotional and cognitive trust in AI negatively correlated with female respondents in 14 out of 16 cases, with significant correlations in 6 of those (38%) (see Appendix 5-12). Additionally, we observed a positive correlation between gender and technology anxiety (TA) in all groups, with significance in 3 out of 4 groups (75%). This suggests that women tend to have lower trust in AI and higher levels of TA, a finding consistent with existing research on gender and anxiety (McLean & Anderson, 2009). However, it is important to note that our sample‘s female majority could have influenced these results, potentially leading to skewed outcomes.[23]

5.3.2. Respondents' rationale

Analyzing comments from respondents provides insights into their reasoning. In the Dating scenario, preferences for humans emerged due to privacy concerns and the need for deeper personal understanding. The Math scenario comments revealed doubts about AI‘s ability to handle nuances, like interpreting poor handwriting or understanding the reasoning behind calculations, areas where human judgment was deemed superior. However, some saw AI as preferable for its consistency, error minimization and speed. In the Doctor scenario, AI‘s consistent, unbiased performance was seen as beneficial, yet many still valued human interaction for this task. The Career scenario highlighted the importance of personal feelings, but also recognized AI‘s cost-effectiveness, benefiting those with financial constraints.

Overall, comments showed a trend: preference for humans in tasks requiring personal recognition, trust and subjective interpretation, preference for AI in tasks demanding efficiency, technical accuracy and minimal errors, and no preference where AI and humans were perceived to complement each other or in highly subjective tasks where neither was seen as fully capable. This respondent rationale aligns with Hypothesis 4, suggesting a positive correlation between the tendency to rely on AI and the perceived objectiveness of the task.

5.3.3. Other insights

Another noteworthy discovery is that overall, a large proportion of answers, nearly one quarter (24%) expressed “no preference” between an AI and a human performing the task in questions (see Figure 7). It is possible that a portion of these respondents might prefer to undertake the task themselves (as seen in comments from respondents), which is supported from the data when looking at the *follow recommendation* score for the doctor and math scenario (two scenarios which are not self-service in nature, i.e. most individuals cannot perform these tasks themselves), where they score the highest.

[23] *T-tests could have been done to compare averages between genders, however due to time constraints as well as the likelihood of skewed results, the decision was made to only briefly comment on the correlation matrices.*

5.4. Suggestions for future research

Future research in the field of AI trust among Swedish adults offers vast potential for deeper insights. Our study, centered on young Swedish adults, predominantly female students, with a restricted sample size, serves as a starting point. Future research should aim for larger, more diverse samples encompassing different ages, genders, cultural backgrounds and AI exposure levels. Drawing inspiration from Glikson & Woolley (2020) emphasis on cultural context, an expanded sample would enable a more detailed understanding of how trust in AI varies across various cultures and age groups.

Five to ten years ago, AI performing tasks like matchmaking in dating apps, grading math exams, diagnosing diseases or offering career advice might have seemed far-fetched. Yet, our study shows these scenarios are now seen as highly realistic. Future research on AI trust in specific tasks should update scenarios to reflect ongoing technological advancements, considering tasks once viewed as unrealistic for AI may become feasible. Moreover, future studies should account for the influence of the current media landscape and the potential temporary impact of trending topics on individuals‘ responses.

As discussed earlier, future research could delve deeper into participants‘ AI interactions, including the context and impact of these interactions (Castelo et al. 2019; Johnson and Grayson, 2005). This might involve directly asking how familiar participants feel with AI (similar to Glikson & Wolley, 2020), then comparing this self-assessment with their actual usage and frequency of AI applications. Longitudinal studies would be particularly effective, tracking how trust in AI evolves over time. Such an approach, in line with Glikson & Woolley (2020) trust trajectory concepts, would provide a dynamic understanding crucial for keeping pace with the rapidly evolving AI landscape.

Future research should also keep up with AI‘s rapid evolution and if necessary, refine how AI-related anxiety is measured. Current methods like the Technological Anxiety (TA) scale are effective and established, but ongoing AI advancements could eventually make them outdated. Studies by Li et al. (2020) and Wang and Wang (2022) represent progress, but further development and establishment is needed.

Finally, future studies could advantageously measure both intention and actual behavior to explore if there are any potential differences between the two as well as investigate how ethical implications and regulations impact trust.

6. REFERENCES

Aditya Malik. (2023). *AI Bias In Recruitment: Ethical Implications And Transparency*. 2023. Retrieved December 5, 2023, from https://www.forbes.com/sites/forbestechcouncil/2023/09/25/ai-bias-in-recruitment-ethical-implications-and-transparency/?sh=4f4b9106799f

Alex Tapscott. (2023). *The Sam Altman saga reveals the need for AI transparency*. https://nypost.com/2023/11/25/opinion/the-sam-altman-saga-reveals-the-need-for-ai-transparency/

Alin, A. (2010). Multicollinearity. *Wiley Interdisciplinary Reviews: Computational Statistics*, *2*(3), 370–374. https://doi.org/10.1002/WICS.84

Antes, A. L., Burrous, S., Sisk, B. A., Schuelke, M. J., Keune, J. D., & DuBois, J. M. (2021). Exploring perceptions of healthcare technologies enabled by artificial intelligence: an online, scenario-based survey. *BMC Medical Informatics and Decision Making*, *21*(1). https://doi.org/10.1186/S12911-021-01586-8

Araujo, T., Helberger, N., Kruikemeier, S., & de Vreese, C. H. (2020). In AI we trust? Perceptions about automated decision-making by artificial intelligence. *AI and Society*, *35*(3), 611–623. https://doi.org/10.1007/S00146-019-00931-W

Asan, O., Bayrak, A. E., & Choudhury, A. (2020). Artificial Intelligence and Human Trust in Healthcare: Focus on Clinicians. In *Journal of Medical Internet Research* (Vol. 22, Issue 6). JMIR Publications Inc. https://doi.org/10.2196/15154

Bell, E., Harley, B., Bryman, A. & B., & Bell, Emma, Alan Bryman, and B. H. (2022). Business research methods. *2022, Sixth Edition.*, 18–48. https://books.google.com/books/about/Business_Research_Methods.html?hl=sv&id=hptjEAAAQBAJ

Ben Mimoun, M. S., Poncin, I., & Garnier, M. (2012). Case study-Embodied virtual agents: An analysis on reasons for failure. *Journal of Retailing and Consumer Services*, *19*(6), 605–612. https://doi.org/10.1016/J.JRETCONSER.2012.07.006

Bhatnagar, S., Alexandrova, A., Avin, S., Cave, S., Cheke, L., Crosby, M., Feyereisl, J., Halina, M., Loe, B. S., Seán´o, S., Eigeartaigh, ´, Martínez-Plumed, F., Price, H., Shevlin, H., Weller, A., Winfield, A., & Hernández-Orallo, J. (2018). *Mapping Intelligence: Requirements and Possibilities*.

Blauth, T. F., Gstrein, O. J., & Zwitter, A. (2022). Artificial Intelligence Crime: An Overview of Malicious Use and Abuse of AI. *IEEE Access*, *10*, 77110–77122. https://doi.org/10.1109/ACCESS.2022.3191790

Brynjolfsson, E., & Mcafee, A. (2011). *Race Against the Machine*. www.RaceAgainstTheMachine.com

Castelo, N., Bos, M. W., & Lehmann, D. R. (2019). Task-Dependent Algorithm Aversion. *Journal of Marketing Research*, *56*(5), 809–825. https://doi.org/10.1177/0022243719851788

Collins, C., Dennehy, D., Conboy, K., & Mikalef, P. (2021). Artificial intelligence in information systems research: A systematic literature review and research agenda. *International Journal of Information Management*, *60*, 102383. https://doi.org/10.1016/J.IJINFOMGT.2021.102383

Därför sågas självkörande taxibilar | Göteborgs-Posten. (2023). Retrieved December 5, 2023, from https://www.gp.se/ekonomi/darfor-sagas-sjalvkorande-taxibilar.080d7285-270a-5f3b-8fc1-3318ee5b1da6

Davis, F. D., Bagozzi, R. P., & Warshaw, P. R. (1989). User Acceptance of Computer Technology: A Comparison of Two Theoretical Models. *Source: Management Science*, *35*(8), 982–1003. https://www.jstor.org/stable/2632151?seq=1&cid=pdf-

Dhami, M. K., Hertwig, R., & Hoffrage, U. (2004). The role of representative design in an ecological approach to cognition. *Psychological Bulletin*, *130*(6), 959–988. https://doi.org/10.1037/0033-2909.130.6.959

Dietvorst, B. J. (2016). *Algorithm aversion*. https://search.proquest.com/openview/b78af8cdf9d7e161ec22d2a9a10e7771/1?pq-origsite=gscholar&cbl=18750

Dietvorst, B. J., Simmons, J. P., & Massey, C. (2018). Overcoming algorithm aversion: People will use imperfect algorithms if they can (even slightly) modify them. *Management Science*, *64*(3), 1155–1170. https://doi.org/10.1287/mnsc.2016.2643

Ertel, W., Black, N., & Mast, F. (2018). *Introduction to artificial intelligence*. 356. https://books.google.com/books/about/Introduction_to_Artificial_Intelligence.html?id=Z4kCuAEACAAJ

Eslami, M., Rickman, A., Vaccaro, K., Aleyasen, A., Vuong, A., Karahalios, K., Hamilton, K., & Sandvig, C. (2015). "I always assumed that I wasn't really that close to [her]": Reasoning about invisible algorithms in news feeds. *Conference on Human Factors in Computing Systems - Proceedings*, *2015-April*, 153–162. https://doi.org/10.1145/2702123.2702556

Ford, M. (Martin R.). (2015). *Rise of the robots : technology and the threat of a jobless future*.

Gaudiello, I., Zibetti, E., Lefort, S., Chetouani, M., & Ivaldi, S. (2016). *Trust as indicator of robot functional and social acceptance. An experimental study on user conformation to iCub answers*. https://doi.org/10.1016/j.chb.2016.03.057ï

Gelbrich, K., & Sattler, B. (2014). Anxiety, crowding, and time pressure in public self-service technology acceptance. *Journal of Services Marketing*, *28*(1), 82–94. https://doi.org/10.1108/JSM-02-2012-0051

Gillath, O., Ai, T., Branicky, M., Keshmiri, S., Davison, R., & Spaulding, R. (2021). Attachment and trust in artificial intelligence. *Computers in Human Behavior*, *115*, 106607. https://doi.org/10.1016/J.CHB.2020.106607

Glikson, E., & Woolley, A. W. (2020). *Human trust in artificial intelligence: Review of empirical research. Academy of Management Annals (in press)*. https://www.researchgate.net/publication/340605601

Groumpos, P. P. (2021). A Critical Historical and Scientific Overview of all Industrial Revolutions. *IFAC-PapersOnLine*, *54*(13), 464–471. https://doi.org/10.1016/J.IFACOL.2021.10.492

Hancock, P. A., Billings, D. R., & Schaefer, K. E. (2011). Can you trust your robot? *Ergonomics in Design*, *19*(3), 24–29. https://doi.org/10.1177/1064804611415045/ASSET/IMAGES/LARGE/10.1177_1064804611415045-FIG3.JPEG

Hengstler, M., Enkel, E., & Duelli, S. (2016). Applied artificial intelligence and trust—The case of autonomous vehicles and medical assistance devices. *Technological Forecasting and Social Change*, *105*, 105–120. https://doi.org/10.1016/J.TECHFORE.2015.12.014

Hoff, K. A., & Bashir, M. (2015). Trust in automation: Integrating empirical evidence on factors that influence trust. *Human Factors*, *57*(3), 407–434. https://doi.org/10.1177/0018720814547570

Hooghe, L., Marks, G., & Schakel, A. H. (2010). The rise of regional authority: A comparative study of 42 democracies. *The Rise of Regional Authority: A Comparative Study of 42 Democracies*, 1–224. https://doi.org/10.4324/9780203852170/RISE-REGIONAL-AUTHORITY-LIESBET-HOOGHE-GARY-MARKS-ARJAN-SCHAKEL

I (Legislative acts) REGULATIONS REGULATION (EU) 2016/679 OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation) (Text with EEA relevance). (2016).

Inbar, Y., Cone, J., & Gilovich, T. (2010). People's intuitions about intuitive insight and intuitive choice. *Journal of Personality and Social Psychology*, *99*(2), 232–247. https://doi.org/10.1037/A0020215

Janiesch, C., Zschech, P., & Heinrich, K. (2021). *Machine learning and deep learning*. https://doi.org/10.1007/s12525-021-00475-2/Published

John Herrman. (2023). *Welcome to the Age of AI-Powered Dating Apps*. https://nymag.com/intelligencer/2023/08/welcome-to-the-age-of-ai-powered-dating-apps.html

John J. Donovan. (1997). *The Second Industrial Revolution: Business Strategy and Internet Technology*. 240.

Johnson, D., & Grayson, K. (2005). Cognitive and affective trust in service relationships. *Journal of Business Research*, *58*(4), 500–507. https://doi.org/10.1016/S0148-2963(03)00140-1

Joseph F. Hair, William C. Black, Barry J. Babin, Ralph E. Anderson, S. (2019). *Multivariate data analysis*. xvii, 813 pages :

https://books.google.com/books/about/Multivariate_Data_Analysis.html?id=0R9ZswEACAAJ

Komiak, S. Y. X., & Benbasat, I. (2006). The effects of personalization and familiarity on trust and adoption of recommendation agents. *MIS Quarterly: Management Information Systems*, *30*(4), 941–960. https://doi.org/10.2307/25148760

Kopp, S., Gesellensetter, L., Krämer, N. C., & Wachsmuth, I. (2005). *A Conversational Agent as Museum Guide-Design and Evaluation of a Real-World Application*.

KPMG. (2023). Trust in Artificial Intelligence: Global Insights 2023 - KPMG Australia. *2023*. https://kpmg.com/au/en/home/insights/2023/02/trust-in-ai-global-insights-2023.html

Krosnick, J. A. (1991). Response strategies for coping with the cognitive demands of attitude measures in surveys. *Applied Cognitive Psychology*, *5*(3), 213–236. https://doi.org/10.1002/ACP.2350050305

Lewandowsky, S., Mundy, M., & Tan, G. P. A. (2000). The dynamics of trust: Comparing humans to automation. *Journal of Experimental Psychology: Applied*, *6*(2), 104–123. https://doi.org/10.1037/1076-898X.6.2.104

Li, J., & Huang, J. S. (2020). Dimensions of artificial intelligence anxiety based on the integrated fear acquisition theory. *Technology in Society*, *63*, 101410. https://doi.org/10.1016/J.TECHSOC.2020.101410

Liu, S. (2012). The relationship between strategic type and new service development competence: a study of Chinese knowledge intensive business services. *Service Business*, *6*(2), 157–175. https://doi.org/10.1007/S11628-011-0122-X

Looije, R., Neerincx, M. A., & Cnossen, F. (2009). Persuasive robotic assistant for health self-management of older adults: Design and evaluation of social behaviors. *Int. J. Human-Computer*. https://doi.org/10.1016/j.ijhcs.2009.08.007

Mayer, R. C., Davis, J. H., & David Schoorman, F. (1995). *An Integrative Model of Organizational Trust* (Vol. 20, Issue 3). https://www.jstor.org/stable/258792?seq=1&cid=pdf-

Mcallister, D. J. (1995). Affect- and Cognition-Based Trust as Foundations for Interpersonal Cooperation in Organizations. In *Source: The Academy of Management Journal* (Vol. 38, Issue 1). https://about.jstor.org/terms

McLean, C. P., & Anderson, E. R. (2009). Brave men and timid women? A review of the gender differences in fear and anxiety. *Clinical Psychology Review*, *29*(6), 496–505. https://doi.org/10.1016/J.CPR.2009.05.003

Meuter, M. L., Ostrom, A. L., Bitner, M. J., & Roundtree, R. (2003). The influence of technology anxiety on consumer use and experiences

with self-service technologies. *Journal of Business Research*, *56*(11), 899–906. https://doi.org/10.1016/S0148-2963(01)00276-4

Mike Isaac, & Cade Metz. (2023). *Meet the A.I. Jane Austen: Meta Weaves A.I. Throughout Its Apps - The New York Times*. 2023. Retrieved December 5, 2023, from https://www.nytimes.com/2023/09/27/technology/meta-ai-celebrities.html

Möhlmann, M., & Zalmanson, L. (2017). *Navigating Algorithmic Management and Drivers*. https://www.researchgate.net/publication/319965259

Musk, Gates och Tallinn varnar: AI hot mot mänskligheten | SvD. (2023). Retrieved February 8, 2024, from https://www.svd.se/a/l3o2kG/musk-gates-och-tallinn-varnar-ai-hot-mot-manskligheten

Oksanen, A., Savela, N., Latikka, R., & Koivula, A. (2020). Trust Toward Robots and Artificial Intelligence: An Experimental Approach to Human–Technology Interactions Online. *Frontiers in Psychology*, *11*. https://doi.org/10.3389/fpsyg.2020.568256

Oppenheimer, D. M., Meyvis, T., & Davidenko, N. (2009). Instructional manipulation checks: Detecting satisficing to increase statistical power. *Journal of Experimental Social Psychology*, *45*(4), 867–872. https://doi.org/10.1016/J.JESP.2009.03.009

O'Shaughnessy, M. R., Schiff, D. S., Varshney, L. R., Rozell, C. J., & Davenport, M. A. (2023). What governs attitudes toward artificial intelligence adoption and governance? *Science and Public Policy*, *50*(2), 161–176. https://doi.org/10.1093/SCIPOL/SCAC056

P3 Nyheter. (2023). *AI-flickvänner kan förstöra en hel generation män 5 oktober 2023 - P3 Tech | Sveriges Radio*. https://sverigesradio.se/avsnitt/ai-flickvanner-kan-forstora-en-hel-generation-man

Ploennigs, J., & Berger, M. (2022). AI Art in Architecture. *AI in Civil Engineering*, *2*(1). https://doi.org/10.1007/s43503-023-00018-y

Rachel Reed. (2023). *AI created a song mimicking the work of Drake and The Weeknd. What does that mean for copyright law? - Harvard Law School | Harvard Law School*. 2023. Retrieved December 5, 2023, from https://hls.harvard.edu/today/ai-created-a-song-mimicking-the-work-of-drake-and-the-weeknd-what-does-that-mean-for-copyright-law/

Raj, M., & Seamans, R. (2019). Primer on artificial intelligence and robotics. *Journal of Organization Design*, *8*(1), 1–14. https://doi.org/10.1186/S41469-019-0050-0/METRICS

Raub, A. (1981). *CORRELATES OF COMPUTER ANXIETY IN COLLEGE STUDENTS*.

Schoorman, F. D., Mayer, R. C., & Davis, J. H. (2007). An Integrative Model of Organizational Trust: Past, Present, and Future. In *Source: The Academy of Management Review* (Vol. 32, Issue 2). https://about.jstor.org/terms

Schwab, K. (2017). *Schwab, K. (2017) The Fourth Industrial Revolution. Crown Publishing Group, New York. - References - Scientific Research Publishing*. https://www.scirp.org/%28S%28351jmbntvnsjt1aadkposzje%29%29/reference/referencespapers.aspx?referenceid=2097936

Scotland, J. (2012). Exploring the philosophical underpinnings of research: Relating ontology and epistemology to the methodology and methods of the scientific, interpretive, and critical research paradigms. *English Language Teaching*, *5*(9), 9–16. https://doi.org/10.5539/ELT.V5N9P9

Statista. (2023). *Artificial Intelligence market size 2030 | Statista*. https://www.statista.com/statistics/1365145/artificial-intelligence-market-size/

Stuart Russell, & Peter Norvig. (1995). *Artificial Intelligence A Modern Approach*.

Svenska Internetstiftelsen. (2023a). *Svenskarna och AI | Svenskarna och internet*. https://svenskarnaochinternet.se/utvalt/svenskarna-och-ai/

Tavakol, M., & Dennick, R. (2011). Making sense of Cronbach's alpha. *International Journal of Medical Education*, *2*, 53. https://doi.org/10.5116/IJME.4DFB.8DFD

Turing, A. M. (2009). Computing machinery and intelligence. *Parsing the Turing Test: Philosophical and Methodological Issues in the Quest for the Thinking Computer*, 23–65. https://doi.org/10.1007/978-1-4020-6710-5_3/COVER

USPTO. (2020). *New benchmark USPTO study finds artificial intelligence in U.S. patents rose by more than 100% since 2002 | USPTO*. https://www.uspto.gov/about-us/news-updates/new-benchmark-uspto-study-finds-artificial-intelligence-us-patents-rose-more

Wang, W., Qiu, L., Kim, D., & Benbasat, I. (2016). Effects of rational and social appeals of online recommendation agents on cognition- and affect-based trust. *Decision Support Systems*, *86*, 48–60. https://doi.org/10.1016/J.DSS.2016.03.007

Wang, Y. Y., & Wang, Y. S. (2022). Development and validation of an artificial intelligence anxiety scale: an initial application in predicting motivated learning behavior. *Interactive Learning Environments*, *30*(4), 619–634. https://doi.org/10.1080/10494820.2019.1674887

World Intellectual Property Organization. (2019). WIPO Technology Trends 2019 – Artificial Intelligence. *WIPO Technology Trends 2019*. https://doi.org/10.34667/TIND.29084

Xu, M., David, J. M., & Kim, S. H. (2018). The Fourth Industrial Revolution: Opportunities and Challenges. *International Journal of Financial Research*, *9*(2), 90–95. https://doi.org/10.5430/IJFR.V9N2P90

Zizzo, D. J. (2011). Experimenter Demand Effects in Economic Experiments. *SSRN Electronic Journal*. https://doi.org/10.2139/SSRN.1163863

7. Appendices

APPENDIX 1. Overview of realism and task characteristics for scenarios in the pilot study

Scenario	Realism	Task characteristics
Dating	4.10	4.85
Career	4.90	4.70
Travel	5.50	4.35
Roman	6.16	4.05
Face cream	5.05	3.50
Painting	5.15	3.60
Invest	5.65	3.15
Legal	5.00	2.95
Party	6.05	3.85
Doctor	5.65	2.75
Math	5.45	2.4

APPENDIX 2. Distribution of collected responses
"Original" includes responses received from social media, Odenplan metro station and the Central station in Stockholm.

	A	B	C	D	Total
Original	9	18	17	14	58
SU	37	32	28	33	130
KTH	1	2	2	2	7
Total	47	52	47	49	195

APPENDIX 3. Overview of removed answers

	Entered survey	Complete answers	Failed GDPR/ initials or controls	Answers used	%
Original	169	74	16	58	30%
SU	267	173	43	130	67%
KTH	25	8	1	7	4%
Total	461	255	60	195	

Note: Original= the stockholm Central station, Odenplan and social media

APPENDIX 4. Cronbach's alpha for items related to emotional trust, cognitive trust and TA

	Emotional trust	Cognitive trust
A1	0.86	0.68
A2	0.88	0.84
B1	0.86	0.66
B2	0.91	0.69
C1	0.76	0.83
C2	0.76	0.87
D1	0.72	0.86
D2	0.77	0.70
TA	0.79	

APPENDIX 5. Correlation matrix A1

	Obj. A1	FR. A1	Avg.EmT. A1	Avg.CogT. A1	Avg. TA	Gender	Familiarity	Initial R	Ini.T. AI	In.T.H
Obj. A1	1	0.11	-0.01	0.21	-0.01	-0.03	0	-0.1	0.13	0.03
FR. A1	0.11	1	0.24	0.43**	-0.27	0.03	0.26	0.3*	-0.07	0.18
Avg.EmT. A1	-0.01	0.24	1	0.65****	-0.68****	-0.32*	0.08	-0.14	0.32*	-0.09
Avg.CogT. A1	0.21	0.43**	0.65****	1	-0.52***	-0.38**	0.18	-0.26	0.42**	-0.05
Avg. TA	-0.01	-0.27	-0.68****	-0.52***	1	0.35*	-0.26	-0.02	-0.27	0.15
Gender	-0.03	0.03	-0.32*	-0.38**	0.35*	1	0.07	0.33*	-0.27	0.14
Familiarity	0	0.26	0.08	0.18	-0.26	0.07	1	0.09	0.13	0.22
Initial R	-0.1	0.3*	-0.14	-0.26	-0.02	0.33*	0.09	1	-0.15	0.07
Ini.T. AI	0.13	-0.07	0.32*	0.42**	-0.27	-0.27	0.13	-0.15	1	-0.2
In.T.H	0.03	0.18	-0.09	-0.05	0.15	0.14	0.22	0.07	-0.2	1

Explanation of abbreviations for the matrices:
Obj.: Objectivism
FR.: Follow Recommendation
Avg.EmT.: Average Emotional Trust
Avg.CogT.: Average Cognitive Trust
Initial R: Initial Regulation
Ini.T. AI: Initial trust in AI
In.T.H: Initial trust in Humans

APPENDIX 6. Correlation matrix A2

	Obj. A2	FR. A2	Avg.EmT. A2	Avg.CogT. A2	Avg. TA	Gender	Familiarity	Initial R	Ini.T. AI	In.T.H
Obj. A2	1	-0.1	-0.29*	-0.41**	0.17	0.06	0.08	-0.28	0.03	0.08
FR. A2	-0.1	1	0.22	0.39**	-0.08	-0.28	-0.23	-0.09	0.45**	-0.02
Avg.EmT. A2	-0.29*	0.22	1	0.65****	-0.58****	-0.12	0.07	0	0.19	-0.14
Avg.CogT. A2	-0.41**	0.39**	0.65****	1	-0.18	-0.04	0.05	0.11	0.26	-0.08
Avg. TA	0.17	-0.08	-0.58****	-0.18	1	0.35*	-0.26	-0.02	-0.27	0.15
Gender	0.06	-0.28	-0.12	-0.04	0.35*	1	0.07	0.33*	-0.27	0.14
Familiarity	0.08	-0.23	0.07	0.05	-0.26	0.07	1	0.09	0.13	0.22
Initial R	-0.28	-0.09	0	0.11	-0.02	0.33*	0.09	1	-0.15	0.07
Ini.T. AI	0.03	0.45**	0.19	0.26	-0.27	-0.27	0.13	-0.15	1	-0.2
In.T.H	0.08	-0.02	-0.14	-0.08	0.15	0.14	0.22	0.07	-0.2	1

APPENDIX 7. Correlation matrix B1

	Obj. B1	FR. B1	Avg.EmT. B1	Avg.CogT. B1	Avg. TA	Gender	Familiarity	Initial R	Ini.T. AI	In.T.H
Obj. B1	1	0.04	0.02	-0.02	0.15	0.18	0	0.01	0.15	0.26
FR. B1	0.04	1	0.01	-0.01	0.04	-0.21	-0.04	0	0.29*	0.2
Avg.EmT. B1	0.02	0.01	1	0.62****	-0.39**	-0.43**	0.21	-0.3*	0.29*	0.09
Avg.CogT. B1	-0.02	-0.01	0.62****	1	-0.37**	-0.23	0.28*	-0.18	0.41**	0.08
Avg. TA	0.15	0.04	-0.39**	-0.37**	1	0.36**	-0.42**	0.18	-0.48***	-0.17
Gender	0.18	-0.21	-0.43**	-0.23	0.36**	1	-0.17	0.06	-0.06	0.02
Familiarity	0	-0.04	0.21	0.28*	-0.42**	-0.17	1	-0.33*	0.5***	0.09
Initial R	0.01	0	-0.3*	-0.18	0.18	0.06	-0.33*	1	-0.24	0.04
Ini.T. AI	0.15	0.29*	0.29*	0.41**	-0.48***	-0.06	0.5***	-0.24	1	0.32*
In.T.H	0.26	0.2	0.09	0.08	-0.17	0.02	0.09	0.04	0.32*	1

APPENDIX 8. Correlation matrix B2

	Obj. B2	FR. B2	Avg.EmT. B2	Avg.CogT. B2	Avg. TA	Gender	Familiarity	Initial R	Ini.T. AI	In.T.H
Obj. B2	1	-0.19	-0.11	-0.13	0.11	0.27	-0.01	-0.35*	-0.03	-0.09
FR. B2	-0.19	1	0.21	0.25	-0.18	-0.45***	0.15	0.02	0.04	0.01
Avg.EmT. B2	-0.11	0.21	1	0.65****	-0.49***	-0.27	0.24	-0.25	0.27	-0.09
Avg.CogT. B2	-0.13	0.25	0.65****	1	-0.38**	-0.27	0.15	-0.08	0.21	0.22
Avg. TA	0.11	-0.18	-0.49***	-0.38**	1	0.36**	-0.42**	0.18	-0.48***	-0.17
Gender	0.27	-0.45***	-0.27	-0.27	0.36**	1	-0.17	0.06	-0.06	0.02
Familiarity	-0.01	0.15	0.24	0.15	-0.42**	-0.17	1	-0.33*	0.5***	0.09
Initial R	-0.35*	0.02	-0.25	-0.08	0.18	0.06	-0.33*	1	-0.24	0.04
Ini.T. AI	-0.03	0.04	0.27	0.21	-0.48***	-0.06	0.5***	-0.24	1	0.32*
In.T.H	-0.09	0.01	-0.09	0.22	-0.17	0.02	0.09	0.04	0.32*	1

APPENDIX 9. Correlation matrix C1

	Obj. C1	FR. C1	Avg.EmT. C1	Avg.CogT. C1	Avg. TA	Gender	Familiarity	Initial R	Ini.T. AI	In.T.H
Obj. C1	1	0.01	-0.08	-0.16	-0.12	-0.17	-0.01	-0.08	-0.25	-0.28
FR. C1	0.01	1	0.11	0.26	-0.03	-0.07	0.18	-0.18	0.19	0.35*
Avg.EmT. C1	-0.08	0.11	1	0.46**	-0.38**	-0.3*	0.2	-0.02	0.01	-0.32*
Avg.CogT. C1	-0.16	0.26	0.46**	1	-0.38**	-0.19	0.29	-0.21	0.42**	0.01
Avg. TA	-0.12	-0.03	-0.38**	-0.38**	1	0.06	-0.59****	0.28	-0.2	-0.11
Gender	-0.17	-0.07	-0.3*	-0.19	0.06	1	-0.13	-0.04	-0.02	0.1
Familiarity	-0.01	0.18	0.2	0.29	-0.59****	-0.13	1	-0.14	0.1	0.14
Initial R	-0.08	-0.18	-0.02	-0.21	0.28	-0.04	-0.14	1	-0.2	0.06
Ini.T. AI	-0.25	0.19	0.01	0.42**	-0.2	-0.02	0.1	-0.2	1	0.4**
In.T.H	-0.28	0.35*	-0.32*	0.01	-0.11	0.1	0.14	0.06	0.4**	1

APPENDIX 10. Correlation matrix C2

	Obj. C2	FR. C2	Avg.EmT. C2	Avg.CogT. C2	Avg. TA	Gender	Familiarity	Initial R	Ini.T. AI	In.T.H
Obj. C2	1	-0.03	-0.08	0.03	0.11	0.01	-0.03	-0.02	-0.04	-0.12
FR. C2	-0.03	1	0.01	0.23	0.01	-0.23	0.13	-0.03	0.33*	0.25
Avg.EmT. C2	-0.08	0.01	1	0.59****	-0.58****	0.01	0.2	-0.2	0.14	0
Avg.CogT. C2	0.03	0.23	0.59****	1	-0.41**	0.15	0.26	-0.1	0.35*	0.22
Avg. TA	0.11	0.01	-0.58****	-0.41**	1	0.06	-0.59****	0.28	-0.2	-0.11
Gender	0.01	-0.23	0.01	0.15	0.06	1	-0.13	-0.04	-0.02	0.1
Familiarity	-0.03	0.13	0.2	0.26	-0.59****	-0.13	1	-0.14	0.1	0.14
Initial R	-0.02	-0.03	-0.2	-0.1	0.28	-0.04	-0.14	1	-0.2	0.06
Ini.T. AI	-0.04	0.33*	0.14	0.35*	-0.2	-0.02	0.1	-0.2	1	0.4**
In.T.H	-0.12	0.25	0	0.22	-0.11	0.1	0.14	0.06	0.4**	1

APPENDIX 11. Correlation matrix D1

	Obj. D1	FR. D1	Avg.EmT. D1	Avg.CogT. D1	Avg. TA	Gender	Familiarity	Initial R	Ini.T. AI	In.T.H
Obj. D1	1	-0.16	-0.24	-0.24	0.25	0.08	-0.08	-0.1	-0.07	-0.02
FR. D1	-0.16	1	0.27	0.46***	-0.03	-0.21	0.18	-0.09	0.24	-0.06
Avg.EmT. D1	-0.24	0.27	1	0.55****	-0.69****	-0.33*	0.41**	-0.22	0.04	-0.26
Avg.CogT. D1	-0.24	0.46***	0.55****	1	-0.38**	-0.37**	0.37**	0.04	0.37**	-0.06
Avg. TA	0.25	-0.03	-0.69****	-0.38**	1	0.46***	-0.57****	0.3*	-0.27	0.07
Gender	0.08	-0.21	-0.33*	-0.37**	0.46***	1	-0.41**	0	-0.23	0
Familiarity	-0.08	0.18	0.41**	0.37**	-0.57****	-0.41**	1	-0.16	0.32*	-0.19
Initial R	-0.1	-0.09	-0.22	0.04	0.3*	0	-0.16	1	-0.16	-0.07
Ini.T. AI	-0.07	0.24	0.04	0.37**	-0.27	-0.23	0.32*	-0.16	1	0.15
In.T.H	-0.02	-0.06	-0.26	-0.06	0.07	0	-0.19	-0.07	0.15	1

APPENDIX 12. Correlation matrix D2

	Obj. D2	FR. D2	Avg.EmT. D2	Avg.CogT. D2	Avg. TA	Gender	Familiarity	Initial R	Ini.T. AI	In.T.H
Obj. D2	1	0	-0.08	-0.19	0.05	-0.1	-0.01	0.28	0.05	-0.2
FR. D2	0	1	0.19	0.38**	-0.13	0.04	0.09	-0.04	0.05	-0.11
Avg.EmT. D2	-0.08	0.19	1	0.44**	-0.36*	-0.24	0.4**	-0.07	0.09	-0.21
Avg.CogT. D2	-0.19	0.38**	0.44**	1	-0.22	-0.1	0.24	0.04	0.17	-0.06
Avg. TA	0.05	-0.13	-0.36*	-0.22	1	0.46***	-0.57****	0.3*	-0.27	0.07
Gender	-0.1	0.04	-0.24	-0.1	0.46***	1	-0.41**	0	-0.23	0
Familiarity	-0.01	0.09	0.4**	0.24	-0.57****	-0.41**	1	-0.16	0.32*	-0.19
Initial R	0.28	-0.04	-0.07	0.04	0.3*	0	-0.16	1	-0.16	-0.07
Ini.T. AI	0.05	0.05	0.09	0.17	-0.27	-0.23	0.32*	-0.16	1	0.15
In.T.H	-0.2	-0.11	-0.21	-0.06	0.07	0	-0.19	-0.07	0.15	1

APPENDIX 13. Correlation matrix in between subjects comparison Group A

	Obj. A1	Obj. A2	FR. A1	FR. A2	Avg.EmT. A1	Avg.EmT. A2	Avg.CogT. A1	Avg.CogT. A	Avg. TA	Gender	Familiarity	Initial R	Ini.T. AI	In.T.H
Obj. A1	1	0.34*	0.11	0.2	-0.01	0.05	0.21	0.09	-0.01	-0.03	0	-0.1	0.13	0.03
Obj. A2	0.34*	1	-0.12	-0.1	-0.09	-0.29*	0.08	-0.41**	0.17	0.06	0.08	-0.28	0.03	0.08
FR. A1	0.11	-0.12	1	-0.1	0.24	0.32*	0.43**	0.23	-0.27	0.03	0.26	0.3*	-0.07	0.18
FR. A2	0.2	-0.1	-0.1	1	0.15	0.22	0.21	0.39**	-0.08	-0.28	-0.23	-0.09	0.45**	-0.02
Avg.EmT. A1	-0.01	-0.09	0.24	0.15	1	0.56****	0.65****	0.21	-0.68****	-0.32*	0.08	-0.14	0.32*	-0.09
Avg.EmT. A2	0.05	-0.29*	0.32*	0.22	0.56****	1	0.4**	0.65****	-0.58****	-0.12	0.07	0	0.19	-0.14
Avg.CogT. A1	0.21	0.08	0.43**	0.21	0.65****	0.4**	1	0.25	-0.52***	-0.38**	0.18	-0.26	0.42**	-0.05
Avg.CogT. A2	0.09	-0.41**	0.23	0.39**	0.21	0.65****	0.25	1	-0.18	-0.04	0.05	0.11	0.26	-0.08
Avg. TA	-0.01	0.17	-0.27	-0.08	-0.68****	-0.58****	-0.52***	-0.18	1	0.35*	-0.26	-0.02	-0.27	0.15
Gender	-0.03	0.06	0.03	-0.28	-0.32*	-0.12	-0.38**	-0.04	0.35*	1	0.07	0.33*	-0.27	0.14
Familiarity	0	0.08	0.26	-0.23	0.08	0.07	0.18	0.05	-0.26	0.07	1	0.09	0.13	0.22
Initial R	-0.1	-0.28	0.3*	-0.09	-0.14	0	-0.26	0.11	-0.02	0.33*	0.09	1	-0.15	0.07
Ini.T. AI	0.13	0.03	-0.07	0.45**	0.32*	0.19	0.42**	0.26	-0.27	-0.27	0.13	-0.15	1	-0.2
In.T.H	0.03	0.08	0.18	-0.02	-0.09	-0.14	-0.05	-0.08	0.15	0.14	0.22	0.07	-0.2	1

APPENDIX 14. Correlation matrix in between subjects comparison Group B

	Obj. B1	Obj. B2	FR. B1	FR. B2	Avg.EmT. B1	Avg.EmT. B2	Avg.CogT. B1	Avg.CogT. B	Avg. TA	Gender	Familiarity	Initial R	Ini.T. AI	In.T.H
Obj. B1	1	0.14	0.04	-0.03	0.02	-0.02	-0.02	-0.08	0.15	0.18	0	0.01	0.15	0.26
Obj. B2	0.14	1	-0.12	-0.19	-0.06	-0.11	0.13	-0.13	0.11	0.27	-0.01	-0.35*	-0.03	-0.09
FR. B1	0.04	-0.12	1	0.25	0.01	0.11	-0.01	0.04	0.04	-0.21	-0.04	0	0.29*	0.2
FR. B2	-0.03	-0.19	0.25	1	-0.05	0.21	0.01	0.25	-0.18	-0.45***	0.15	0.02	0.04	0.01
Avg.EmT. B1	0.02	-0.06	0.01	-0.05	1	0.42**	0.62****	0.23	-0.39**	-0.43**	0.21	-0.3*	0.29*	0.09
Avg.EmT. B2	-0.02	-0.11	0.11	0.21	0.42**	1	0.26	0.65****	-0.49***	-0.27	0.24	-0.25	0.27	-0.09
Avg.CogT. B1	-0.02	0.13	-0.01	0.01	0.62****	0.26	1	0.26	-0.37**	-0.23	0.28*	-0.18	0.41**	0.08
Avg.CogT. B2	-0.08	-0.13	0.04	0.25	0.23	0.65****	0.26	1	-0.38**	-0.27	0.15	-0.08	0.21	0.22
Avg. TA	0.15	0.11	0.04	-0.18	-0.39**	-0.49***	-0.37**	-0.38**	1	0.36**	-0.42**	0.18	-0.48***	-0.17
Gender	0.18	0.27	-0.21	-0.45***	-0.43**	-0.27	-0.23	-0.27	0.36**	1	-0.17	0.06	-0.06	0.02
Familiarity	0	-0.01	-0.04	0.15	0.21	0.24	0.28*	0.15	-0.42**	-0.17	1	-0.33*	0.5***	0.09
Initial R	0.01	-0.35*	0	0.02	-0.3*	-0.25	-0.18	-0.08	0.18	0.06	-0.33*	1	-0.24	0.04
Ini.T. AI	0.15	-0.03	0.29*	0.04	0.29*	0.27	0.41**	0.21	-0.48***	-0.06	0.5***	-0.24	1	0.32*
In.T.H	0.26	-0.09	0.2	0.01	0.09	-0.09	0.08	0.22	-0.17	0.02	0.09	0.04	0.32*	1

APPENDIX 15. Correlation matrix in between subjects comparison Group C

	Obj. C1	Obj. C2	FR. C1	FR. C2	Avg.EmT. C1	Avg.EmT. C2	Avg.CogT. C1	Avg.CogT. C2	Avg. TA	Gender	Familiarity	Initial I	Ini.T. AI	In.T.H
Obj. C1	1	0.06	0.01	-0.02	-0.08	0.15	-0.16	0.11	-0.12	-0.17	-0.01	-0.08	-0.25	-0.28
Obj. C2	0.06	1	0.34*	-0.03	0.14	-0.08	0.15	0.03	0.11	0.01	-0.03	-0.02	-0.04	-0.12
FR. C1	0.01	0.34*	1	0.41**	0.11	0.15	0.26	0.3*	-0.03	-0.07	0.18	-0.18	0.19	0.35*
FR. C2	-0.02	-0.03	0.41**	1	-0.03	0.01	0.39**	0.23	0.01	-0.23	0.13	-0.03	0.33*	0.25
Avg.EmT. C1	-0.08	0.14	0.11	-0.03	1	0.33*	0.46**	0.1	-0.38**	-0.3*	0.2	-0.02	0.01	-0.32*
Avg.EmT. C2	0.15	-0.08	0.15	0.01	0.33*	1	0.07	0.59****	-0.58****	0.01	0.2	-0.2	0.14	0
Avg.CogT. C1	-0.16	0.15	0.26	0.39**	0.46**	0.07	1	0.3*	-0.38**	-0.19	0.29	-0.21	0.42**	0.01
Avg.CogT. C2	0.11	0.03	0.3*	0.23	0.1	0.59****	0.3*	1	-0.41**	0.15	0.26	-0.1	0.35*	0.22
Avg. TA	-0.12	0.11	-0.03	0.01	-0.38**	-0.58****	-0.38**	-0.41**	1	0.06	-0.59****	0.28	-0.2	-0.11
Gender	-0.17	0.01	-0.07	-0.23	-0.3*	0.01	-0.19	0.15	0.06	1	-0.13	-0.04	-0.02	0.1
Familiarity	-0.01	-0.03	0.18	0.13	0.2	0.2	0.29	0.26	-0.59****	-0.13	1	-0.14	0.1	0.14
Initial R	-0.08	-0.02	-0.18	-0.03	-0.02	-0.2	-0.21	-0.1	0.28	-0.04	-0.14	1	-0.2	0.06
Ini.T. AI	-0.25	-0.04	0.19	0.33*	0.01	0.14	0.42**	0.35*	-0.2	-0.02	0.1	-0.2	1	0.4**
In.T.H	-0.28	-0.12	0.35*	0.25	-0.32*	0	0.01	0.22	-0.11	0.1	0.14	0.06	0.4**	1

APPENDIX 16. Correlation matrix in between subjects comparison Group D

	Obj. D1	Obj. D2	FR. D1	FR. D2	Avg.EmT. D1	Avg.EmT. D2	Avg.CogT. D1	Avg.CogT. D2	Avg. TA	Gender	Familiarity	Initial I	Ini.T. AI	In.T.H
Obj. D1	1	0.14	-0.16	-0.04	-0.24	-0.16	-0.24	-0.05	0.25	0.08	-0.08	-0.1	-0.07	-0.02
Obj. D2	0.14	1	0.2	0	0.04	-0.08	-0.02	-0.19	0.05	-0.1	-0.01	0.28	0.05	-0.2
FR. D1	-0.16	0.2	1	0.13	0.27	0.12	0.46***	-0.05	-0.03	-0.21	0.18	-0.09	0.24	-0.06
FR. D2	-0.04	0	0.13	1	0.01	0.19	-0.08	0.38**	-0.13	0.04	0.09	-0.04	0.05	-0.11
Avg.EmT. D1	-0.24	0.04	0.27	0.01	1	0.52***	0.55****	0.26	-0.69****	-0.33*	0.41**	-0.22	0.04	-0.26
Avg.EmT. D2	-0.16	-0.08	0.12	0.19	0.52***	1	0.35*	0.44**	-0.36*	-0.24	0.4**	-0.07	0.09	-0.21
Avg.CogT. D1	-0.24	-0.02	0.46***	-0.08	0.55****	0.35*	1	0.36*	-0.38**	-0.37**	0.37**	0.04	0.37**	-0.06
Avg.CogT. D2	-0.05	-0.19	-0.05	0.38**	0.26	0.44**	0.36*	1	-0.22	-0.1	0.24	0.04	0.17	-0.06
Avg. TA	0.25	0.05	-0.03	-0.13	-0.69****	-0.36*	-0.38**	-0.22	1	0.46***	-0.57****	0.3*	-0.27	0.07
Gender	0.08	-0.1	-0.21	0.04	-0.33*	-0.24	-0.37**	-0.1	0.46***	1	-0.41**	0	-0.23	0
Familiarity	-0.08	-0.01	0.18	0.09	0.41**	0.4**	0.37**	0.24	-0.57****	-0.41**	1	-0.16	0.32*	-0.19
Initial R	-0.1	0.28	-0.09	-0.04	-0.22	-0.07	0.04	0.04	0.3*	0	-0.16	1	-0.16	-0.07
Ini.T. AI	-0.07	0.05	0.24	0.05	0.04	0.09	0.37**	0.17	-0.27	-0.23	0.32*	-0.16	1	0.15
In.T.H	-0.02	-0.2	-0.06	-0.11	-0.26	-0.21	-0.06	-0.06	0.07	0	-0.19	-0.07	0.15	1

APPENDIX 17. Regression analysis on emotional trust

Dependent variable - Emotional trust

	Dating	Math	Doctor	Career
AI/H	-0.226 (0.223)	0.247 (0.188)	-0.408** (0.168)	-0.111 (0.174)
Task characteristics	0.031 (0.068)	-0.084 (0.062)	-0.068 (0.058)	-0.028 (0.052)
Experience	-0.031 (0.060)	-0.018 (0.059)	-0.009 (0.053)	0.037 (0.054)
TA	-0.520*** (0.087)	-0.487*** (0.086)	-0.461*** (0.091)	-0.318*** (0.093)
Constant	4.718*** (0.488)	5.320*** (0.436)	5.003*** (0.434)	4.562*** (0.448)
Observations	99	99	96	96
R2	0.292	0.309	0.337	0.200
Adjusted R2	0.262	0.280	0.308	0.165
Residual Std. Error	0.914 (df = 94)	0.894 (df = 94)	0.799 (df = 91)	0.814 (df = 91)
F Statistic	9.700*** (df = 4;94)	10.525*** (df = 4;94)	11.588*** (df = 4;91)	5.682*** (df = 4;91)
Note:	*p<0.1;	**p<0.05;	***p<0.01	

APPENDIX 18. Regression analysis on cognitive trust

Dependent variable - Cognitive trust

	Dating	Math	Doctor	Career
AI/H	-0.109 (0.157)	0.411*** (0.153)	-0.277* (0.149)	0.043 (0.159)
Task characteristics	0.064 (0.048)	-0.122** (0.050)	-0.089* (0.052)	-0.027 (0.047)
Experience	0.035 (0.042)	0.009 (0.048)	0.072 (0.047)	0.052 (0.050)
TA	-0.267*** (0.061)	-0.156** (0.070)	-0.179** (0.081)	-0.154* (0.085)
Constant	3.968*** (0.345)	4.579*** (0.354)	4.689*** (0.385)	4.183*** (0.409)
Observations	99	99	96	96
R2	0.218	0.168	0.207	0.100
Adjusted R2	0.185	0.133	0.173	0.061
Residual Std. Error	0.645 (df = 94)	0.725 (df = 94)	0.708 (df = 91)	0.744 (df = 91)
F Statistic	6.550*** (df = 4; 94)	4.751*** (df = 4; 94)	5.955*** (df = 4; 91)	5.530*** (df = 4; 91)
Note:	*p<0.1;	**p<0.05;	***p<0.01	

APPENDIX 19. Regression analysis on follow recommendation

	Dependent variable - Follow recommendation			
	Dating	**Math**	**Doctor**	**Career**
AI/H	0.480 (0.360)	1.104*** (0.389)	0.177 (0.379)	0.411 (0.280)
Task characteristics	0.089 (0.110)	-0.158 (0.128)	-0.075 (0.131)	-0.010 (0.083)
Experience	0.072 (0.096)	-0.110 (0.121)	0.226* (0.120)	0.074 (0.087)
TA	-0.102 (0.140)	-0.237 (0.178)	0.175 (0.206)	-0.011 (0.149)
Constant	3.232*** (0.788)	5.880*** (0.902)	3.634*** (0.978)	3.809*** (0.720)
Observations	99	99	96	96
R2	0.068	0.112	0.048	0.042
Adjusted R2	0.029	0.074	0.006	-0.0003
Residual Std. Error	1.474 (df = 94)	1.850 (df = 94)	1.799 (df = 91)	1.308 (df = 91)
F Statistic	1.725 (df = 4; 94)	2.954** (df = 4; 94)	1.137 (df = 4; 94)	0.992 (df = 4; 94)
Note:	*p<0.1;	**p<0.05;	***p<0.01	

APPENDIX 20. Variance Inflation Factors

VIF emotional trust

	AI/H	Task characteristics	Experience	TA
Dating	1,47	1,46	1,13	1,14
Math	1,09	1,10	1,14	1,16
Doctor	1,06	1,03	1,49	1,46
Career	1,10	1,06	1,49	1,46

VIF cognitive trust

	AI/H	Task characteristics	Experience	TA
Dating	1,47	1,46	1,13	1,14
Math	1,09	1,10	1,14	1,16
Doctor	1,06	1,03	1,49	1,46
Career	1,10	1,06	1,49	1,46

VIF follow recommendation

	AI/H	Task characteristics	Experience	TA
Dating	1,47	1,46	1,13	1,14
Math	1,09	1,10	1,14	1,16
Doctor	1,06	1,03	1,49	1,46
Career	1,10	1,06	1,49	1,46

APPENDIX 21. T-test values per scenario

AI vs Human scenario

	Dating	**Math**	**Doctor**	**Career**
Variables	A1 vs B1	A2 vs B2	C1 vs D1	C2 vs D2
Realism	0.994	**<0.001**	0.270	0.950
Task characteristics	**<0.001**	**<0.001**	**<0.001**	0.014
Follow recommendation	0.032	**0.008**	0.446	0.095
Emotional trust	0.644	0.257	**0.02**	0.515
Cognitive trust	0.812	**0.037**	**0.084**	0.753

***Note**: Bold numbers are significant*

APPENDIX 22. T-test values subjective vs objective, emotional vs cognitive trust

Emotional vs. Cognitive trust

A1 vs A1	A2 vs A2	B1 vs B1	B2 vs B2	C1 vs C1	C2 vs C2	D1 vs D1	D2 vs D2
0.037	0.159	**0.002**	**0.025**	**<0.001**	0.089	**<0.001**	0.408

***Note** : Bold numbers are significant*

Subjective vs. objective	**A**	**B**	**C**	**D**
	A1 vs A2	B1 vs B2	C1 vs C2	D1 vs D2
Emotional trust	0.058	**<0.001**	0.06	**0.022**
Cognitive trust	0.056	**<0.001**	0.934	**0.020**

***Note** : Bold numbers are significant*

APPENDIX 23. Post-hoc Tukey test

Variable	***Group comparison***	**Mean Difference**	**Lower bound**	**Upper bound**	**Pr(>F**
Task characteristics	Group A and B	1.32	0.71	1.92	<0.001 **
Task characteristics	Group C and D	(0.60)	(1.21)	0.01	0.055
Follow recommendation	Group A and B	0.83	0.22	1.45	<0.005 **
Follow recommendation	Group C and D	(0.37)	(0.99)	0.26	0.429
Cognitive trust	Group A and B	0.18	(0.10)	0.47	0.357
Cognitive trust	Group C and D	0.11	0.18	0.40	0.745

***Note** : ** Statistically significant since p<0.05*

APPENDIX 24. ANOVA for experience, regulation, general trust AI

Variable	***SumSq***	***Mean Sq***	***F-value***	***Pr(>F***
Experience	102.8	3.388	1.093	0.353
Regulation	43.3	1.017	1.475	0.223
General Trust AI	4.9	0.748	1.122	0.341
General Trust Human	5.36	0.0763	0.171	0.916

APPENDIX 25. Scenarios (translated from Swedish to English):

Dating

A1: Imagine you‘re single and looking for a romantic partner. A dating app has developed an AI that provides personalized recommendations for potential partners that match your preferences. This AI has successfully matched other friends in the past.

B1: Imagine you‘re single and looking for a romantic partner. A friend wants to set you up with potential partners who match your preferences. This person has successfully matched other friends in the past.

Math exam

A2: Imagine you‘ve just stepped out of the room where you took a university-level mathematics exam that you‘ve been studying intensely for the past few weeks. The exam will be graded by an AI. It is widely known that this AI makes few mistakes in grading.

B2: Imagine you‘ve just stepped out of the room where you took a university-level mathematics exam that you‘ve been studying intensely for the past few weeks. The exam will be graded by your math teacher. It is widely known that this person makes few mistakes in grading.

Doctor

C1: Imagine you have identified a mole on your back and are concerned it might be a cancerous mole. A healthcare company has doctors specialized in diagnosing skin changes, trained to carefully analyze and identify whether a mole is potentially dangerous. These doctors have been recognized multiple times as experts in early detection and treatment of skin-related diseases.

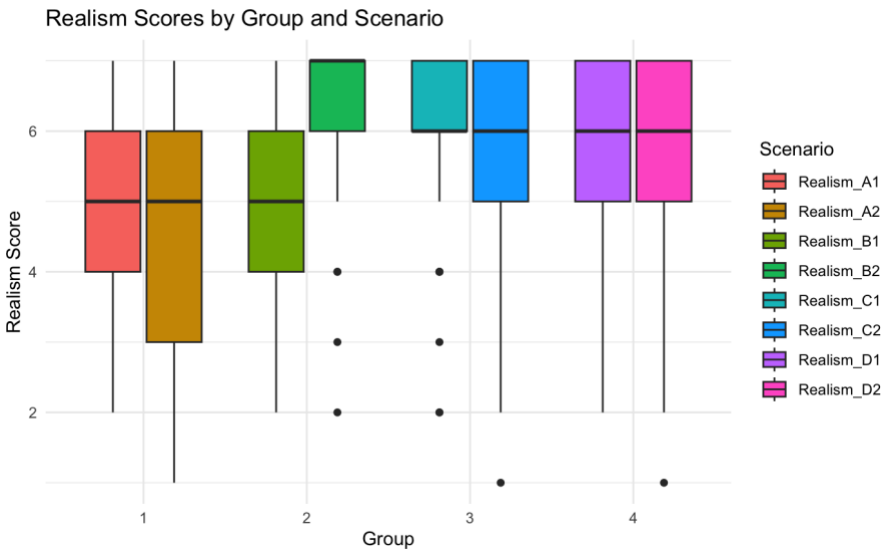
D1: Imagine you have identified a mole on your back and are worried it might be a cancerous mole. A healthcare company has introduced an AI specialized in diagnosing skin changes, trained to carefully analyze and identify whether a mole is potentially dangerous. This AI has been recognized multiple times as an expert in early detection and treatment of skin-related diseases.

Career

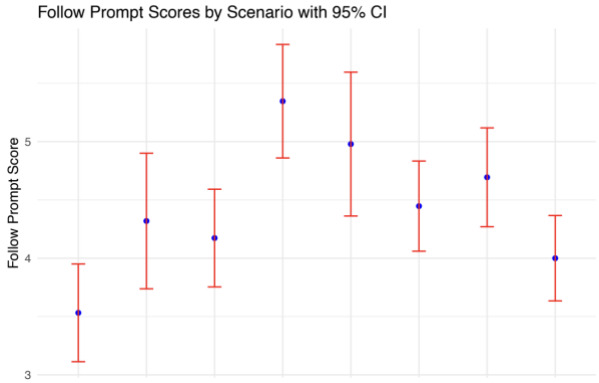
C2: Imagine you‘ve decided to switch career paths. You‘ve heard about a career advisor who analyzes your interests, strengths, weaknesses and goals to recommend the most suitable educational programs and career paths for you. This individual has received excellent reviews.

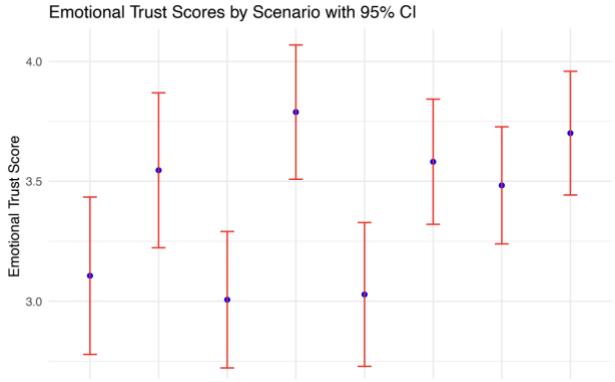
D2: Imagine you‘ve decided to switch career paths. You‘ve heard about an AI-based career advisor that analyzes your interests, strengths, weaknesses and goals to recommend the most suitable educational programs and career paths for you. This AI has received excellent reviews.

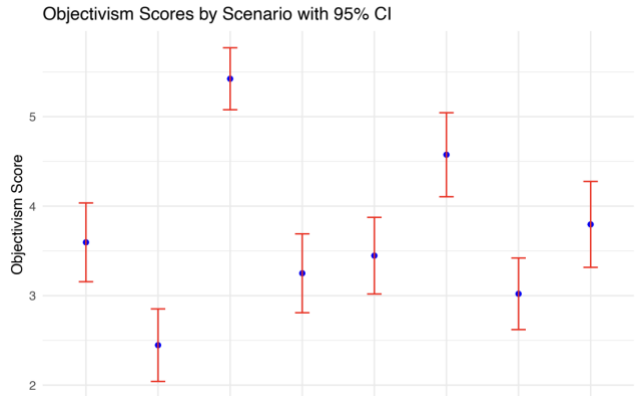
APPENDIX 26. Boxplots for each dependent variable

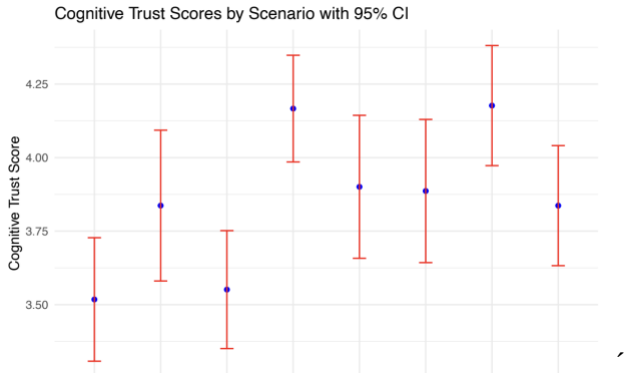


APPENDIX 27. Error bars for relevant variables, showcased on a scenario level from A1 to D2 in following order.

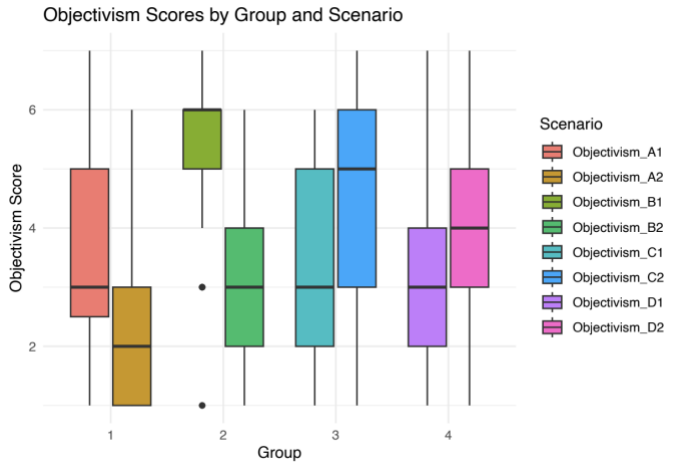


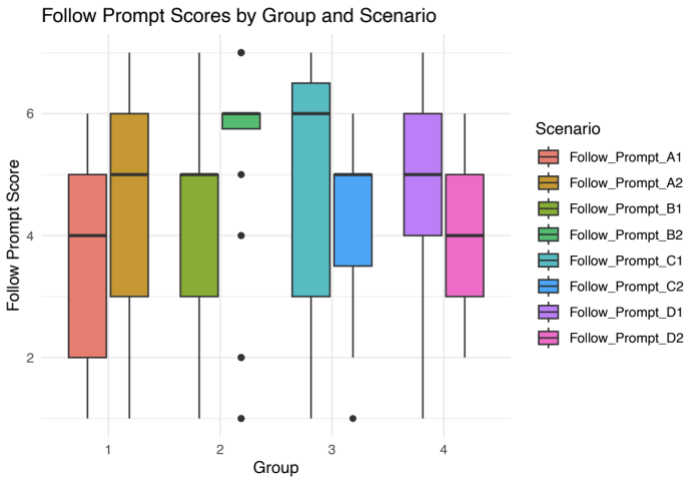






APPENDIX. 28 Box bars for relevant variables, showcased on a scenario level from A1 to D2 in following order.





APPENDIX 29. Questions for emotional and cognitive trust in AI (translated from swedish)

Emotional trust was measured by the following questions; "AI's that can perform this task better than humans makes me uncomfortable", "AI's that can perform this task goes against what I believe technology should be used for" and "AI's capable of performing this type of task are unsettling" and for cognitive trust; "I see the benefits of AI's that can perform this type of task better than humans", "AI's capable of performing this type of task can be useful" and "I believe this type of AI's can perform well".

APPENDIX 30. Main survey

Block 1: Intro
Människa vs. AI 🧠🤖: Vem kan man tro på? 🤔💻

Varmt välkommen till vår undersökning! 🤓 Vi genomför en studie som en del av vår kandidatuppsats vid Handelshögskolan i Stockholm som fokuserar på att utforska hur olika faktorer påverkar förtroendet för AI.💡Vi bjuder in dig till en interaktiv resa där du kommer få utvärdera 2 korta situationer följt av några fördjupande följdfrågor (tar ca 5-8 minuter).

För varje svar donerar vi 2kr till Barncancerfonden! 💌

Om du har frågor om vår studie går det bra att kontakta Rebecka Berg på: 25331@student.hhs.se

Block 2: GDPR
GDPR
Innan vi sätter igång, vänligen läs följande information relaterat till dataskyddsförordningen GDPR.

Projekt: BSc thesis in Business & Economics

År och termin: 2023, höstterminen

Ansvariga studenter för studien: Rebecka Berg, BSc-student (25331@student.hhs.se) samt Gustav Linder, BSc-student (25285student.hhs.se)

Handledare och avdelning vid SSE: Patric Andersson, Associate Professor; Institutionen för marknadsföring och strategi

Handledarens e-postadress: patric.andersson@hhs.se

Typ av personuppgifter om dig som ska behandlas: kön, ålder, utbildningsnivå och sysselsättning.

Information relaterat till GDPR. Som en integrerad del av utbildningsprogrammet vid Handelshögskolan i Stockholm gör inskrivna studenter ett individuellt examensarbete. Detta arbete baseras ibland på undersökningar och intervjuer kopplade till ämnet. Deltagande är naturligtvis helt frivilligt och denna text är avsedd att ge dig nödvändig information om som kan röra ditt deltagande i studien eller intervjun. Du kan när som helst återkalla ditt samtycke och dina uppgifter kommer därefter att raderas permanent.

Sekretess. Allt du säger eller anger i undersökningen eller till intervjuarna kommer att hållas strikt konfidentiellt och kommer endast att göras tillgängligt för handledare, handledare och kursledningsgruppen. Säker lagring av data. All data kommer att lagras och bearbetas säkert av SSE och kommer att raderas permanent när det projekterade är slutfört.

Inga personuppgifter kommer att publiceras. Uppsatsen som skrivs av studenterna kommer inte att innehålla någon information som kan identifiera dig som deltagare i undersökningen eller intervjuämnet.

Dina rättigheter enligt GDPR. Du är välkommen att besöka https://www.hhs.se/en/about- us/data-protection/ för att läsa mer och få information om dina rättigheter relaterade till personuppgifter.

Tveka inte att kontakta oss via mailen; 25331@student.hhs.se om du har frågor kring hur vi hanterar datan!

Block 3: Description AI

Eftersom erfarenheten och kunskapen om AI kan variera mellan deltagare, inleder vi med att först ge en kort allmän förklaring av vad AI är: Artificiell intelligens (AI) är förmågan hos datorprogram/robotar att efterlikna människans naturliga intelligens. Detta inkluderar förmågan att lära sig saker av tidigare erfarenheter, förstå naturligt språk, lösa problem, planera en sekvens av handlingar och att generalisera.

Nedan följer ett antal AI-tjänster, kryssa i dem du någon gång har använt:

- ☐ Large Language Models (ex ChatGPT, Claude, Google Bard, Bing AI)
- ☐ AI-bildgenerering
- ☐ AI-musikgenerering
- ☐ Alexa/Siri/Google assistent
- ☐ Röstgenerering till olika språk/översättning av podcasts och videos i realtid
- ☐ Självkörande/assisterande fordon (ex släppa-ratten funktion)
- ☐ Sociala medier (ex TikTok, Instagram, Facebook)
- ☐ Fitness trackers (ex Fitbit, Apple Watch, Whoop)
- ☐ Smart-home enheter (ex smart högtalare, termostater, madrasser, vitvaror, robotdammsugare mm)

Fyll i hur du ställer dig till följande påstående:

Jag tycker att AI bör regleras.

Stämmer inte alls	Stämmer inte	Varken eller	Stämmer delvis	Stämmer helt

Block 4: General trust

Vad är ditt förtroende generellt sett för uppgifter utförda av AI?

Mycket låg	Ganska låg	Neutral	Ganska hög	Mycket hög

Vad är ditt förtroende generellt sett för uppgifter utförda av människor?

Mycket låg	Ganska låg	Neutral	Ganska hög	Mycket hög

Block 5: Scenario 1 (Dating or Doctor)

See scenarios in Appendix 25

Hur realistisk anser du denna situation vara?

Mycket orealistiskt: Jag kan absolut inte föreställa mig att detta skulle hända i verkligheten.

Ganska orealistiskt: Jag kan inte föreställa mig att detta skulle hända i verkligheten.

Något orealistiskt: Detta känns ganska osannolikt.

Varken eller: Jag är osäker, det kan vara lika sannolikt som osannolikt.

Något realistiskt: Detta känns ganska sannolikt.

Ganska realistiskt: Jag kan föreställa mig att detta skulle hända i verkligheten.

Mycket realistiskt: Jag kan absolut föreställa mig att detta skulle hända i verkligheten.

Hur objektiv/subjektiv anser du denna situation vara?

- Mycket objektiv: Denna situation baseras helt på objektiva uppgifter som är klart definierade och mätbara.
- Ganska objektiv: Uppgiften i denna situation är huvudsakligen kvantifierbar.
- Något objektiv: Denna situation bygger främst på kvantifierbara uppgifter med lite utrymme för personlig tolkning.
- Varken subjektivt eller objektivt: Denna situation har lika mycket kvantifierbara uppgifter som tolkningsutrymme.
- Något subjektiv: Denna situation bygger främst på personlig tolkning med lite utrymme för kvantifierbara uppgifter.
- Ganska subjektiv: Uppgiften i denna situation är huvudsakligen tolkningsbar.
- Mycket subjektiv: Denna situation baseras helt på subjektiva uppgifter så som personliga åsikter, tolkningar och/eller intuition.

Hur troligt är det att du skulle följa AI:s rekommendation för en romantisk partner?

Mycket otroligt	Ganska otroligt	Något otroligt	Varken eller	Något troligt	Ganska troligt	Mycket troligt

Note: This question was adapted in its formulation for each scenario

Nedanstående påståenden handlar om AI:s färdigheter, vänligen ange hur du ställer dig till dessa.

	Stämmer inte alls	Stämmer inte	Varken eller	Stämmer delvis	Stämmer helt
AI:s som kan utföra denna uppgift bättre än människor gör mig obekväm.	○	○	○	○	○
AI:s som kan utföra denna uppgift går emot vad jag tycker teknologi bör användas till.	○	○	○	○	○
AI:s som kan utföra den här typen av uppgift är oroande.	○	○	○	○	○
Jag ser fördelarna med AI:s som kan utföra den här typen av uppgift bättre än människor.	○	○	○	○	○
AI:s som kan utföra den här typen av uppgift kan vara användbara.	○	○	○	○	○
Jag tror att denna typ av AI kan prestera bra.	○	○	○	○	○
På detta påståendet är det viktigt att du väljer “Stämmer inte alls”.	○	○	○	○	○

Skulle du föredra en AI eller en väldigt kvalificerad människa att utföra den här uppgiften?

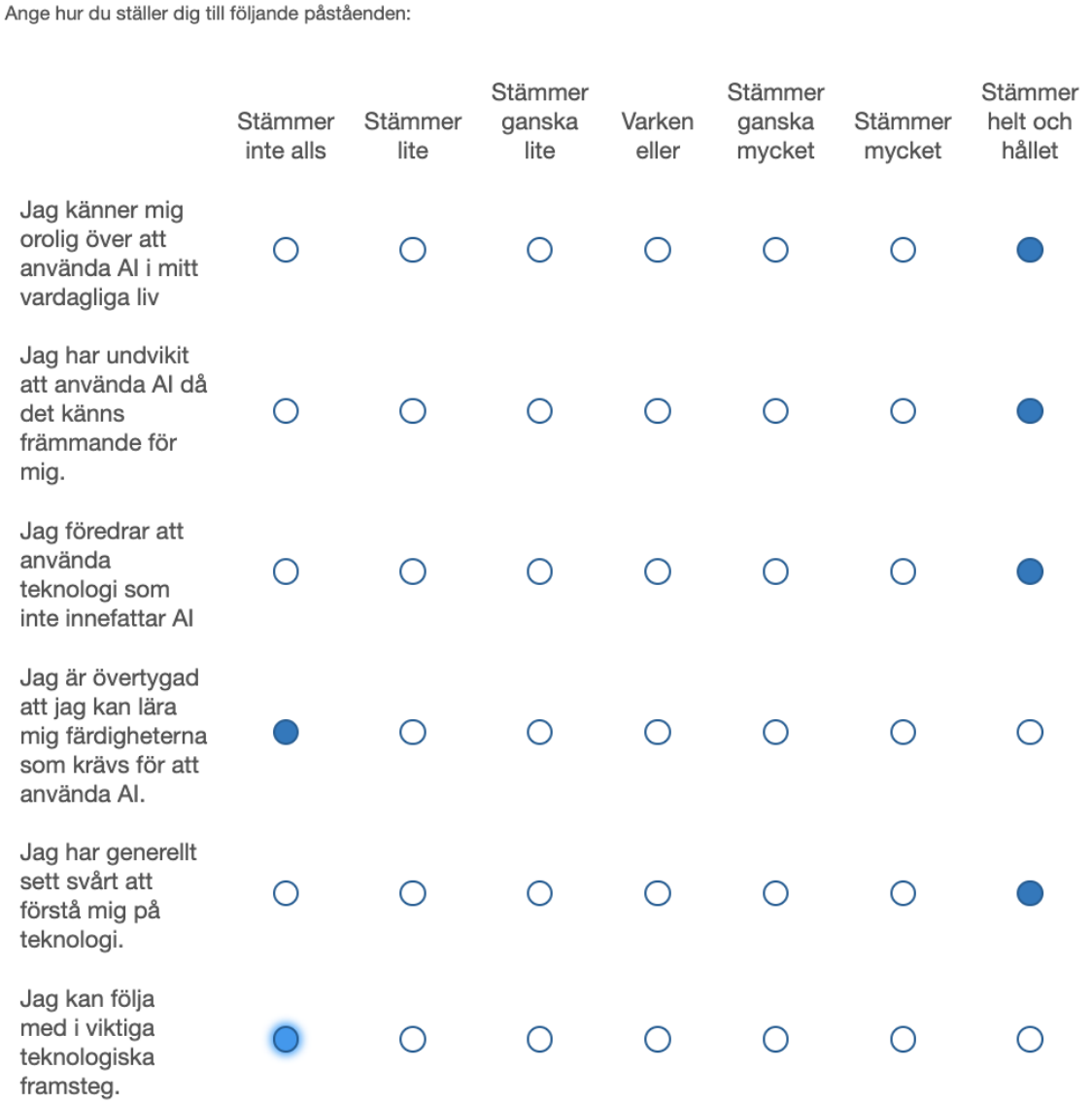
Jag föredrar en människa	Jag föredrar AI	Har ingen åsikt

Plats för kommentar:

Block 6: Scenario 2 (Math or Career)

Questions were the same as for scenario 1 (except the attention check question).

Block 7: TA



Note: The blue dots showcase how question 4 and 6 were reversed as discussed in the Methodology.

Block 8: Demographics

Tack för att du tog dig tid att svara på dessa scenarion! 🙏 Som ett avslutande steg ber vi dig lämna några grundläggande personliga uppgifter om dig själv. Vi följer praxis för dataskydd och integritet, vilket innebär att all information du lämnar kommer att behandlas konfidentiellt och anonymt. 🔒 Dina svar kommer endast att användas i sammanställt skick och kommer inte att kunna kopplas till dig personligen.

Vad är din ålder?

Vad är ditt kön?

Man	Kvinna	Icke-binär/tredje kön	Föredrar att inte svara

Vilken är din utbildningsnivå?

Gymnasium	Högskola/Universitet (ej examen)	Kandidatexamen	Masterexamen eller högre

Vad är din sysselsättning?

Student	Hel/Deltidsanställd	Egen företagare	Arbetslös	Pensionär

Block 9: Quality control questions and attention check

Avslutande kvalitetsfrågor

	Stämmer inte alls	Stämmer inte	Varken eller	Stämmer delvis	Stämmer helt
Jag förstod syftet med denna studie.	○	○	○	○	○
På detta påståendet är det viktigt att du väljer "Stämmer delvis".	○	○	○	○	○
Frågorna i studien var tydligt formulerade.	○	○	○	○	○
Frågorna i studien kändes relevanta för ämnet.	○	○	○	○	○

Appendix 31: Disclosure of AI-tools use

1. What AI tools have been used and how?
In our thesis, we used ChatGPT for various purposes. It was used by interacting with its chatbot on the desktop and in its mobile app.

2. In what ways have these tools contributed to increasing the quality of the thesis?
ChatGPT was used for creative purposes, especially in the ideation phase of our thesis. ChatGPT sped up this process significantly. This increased the speed we were able to move forward, which did not directly affect the quality of the work, but it gave us more time to focus on improving other areas, thereby indirectly contributing to increased quality of our thesis. Moreover, ChatGPT assisted in refining survey questions to be as well formulated as possible. This led to more precise and well-thought-out data from respondents, thus directly affecting the actual quality of the research outcomes. ChatGPT also helped us in debugging in R. This allowed us to create illustrations faster and do more extensive analysis.

3. What potential risks were found using AI and what measures were taken to reduce these risks?
A potential risk could be over-reliance on ChatGPT's suggestions, to mitigate this, we used ChatGPT's inputs to explore and be creative and as a form of guidance, ensuring that the final decisions and interpretations were made independently by us.

4. What are the insights gained from using AI tools in the thesis writing process?
Our experience with using ChatGPT revealed a significant impact on the efficiency of our research process, speeding up various stages of our thesis work. This increased efficiency allowed us to allocate more time to other critical aspects of our research, thereby streamlining our workflow.

However, it's important to note that while ChatGPT enhanced our work process, the influence on the final quality of the thesis was more indirect. ChatGPT served as a support tool rather than a direct contributor to the content of the thesis. This highlights the current role of AI in academic research as a supportive tool that augments rather than replaces us humans. It shows the importance of maintaining a balance between leveraging AI for efficiency and ensuring that the core research and work are driven by humans.