6-3-2025

Moderating Role of Locus of Control on the Relationship Between Self-Efficacy and Trust in AI

Amelia Gillies
Walden University

Walden University

College of Psychology and Community Services

This is to certify that the doctoral dissertation by

Amelia Gillies

has been found to be complete and satisfactory in all respects,
and that any and all revisions required by
the review committee have been made.

Review Committee
Dr. Hedy Dexter, Committee Chairperson, Psychology Faculty
Dr. Anthony Perry, Committee Member, Psychology Faculty

Chief Academic Officer and Provost
Sue Subocz, Ph.D.

Walden University
2025

Abstract

Moderating Role of Locus of Control on the Relationship Between Self-Efficacy and Trust in AI

by

Amelia Gillies

MPhil, Walden University, 2023

PGDipEd, Wesley Institute (Excelsia College), 2010

BA, Monash University, 2006

Dissertation Submitted in Partial Fulfillment

of the Requirements for the Degree of

Doctor of Philosophy

Psychology

Walden University

May 2025

Abstract

The rapid integration of artificial intelligence (AI) into essential sectors, including healthcare, customer services, and defense, has underscored the importance of user trust in AI systems. Psychological traits such as locus of control (the extent to which individuals attribute outcomes to internal or external factors) and self-efficacy (individuals' confidence in their ability) have been identified as significant factors influencing trust in these automated systems. Previous studies have independently linked these psychological traits to trust in technology; however, there has been limited exploration of potential moderating effects of locus of control on the relationship between digital technology self-efficacy and trust in AI. Grounded in Bandura's self-efficacy theory and Rotter's locus of control framework, this quantitative study examined the extent to which locus of control moderated the relationship between digital technology self-efficacy and trust in AI among English-speaking adults, 22 to 55, living in the United States. Using a cross-sectional correlational design, data were collected via Amazon's Mechanical Turk from 125 participants. A moderated multiple regression analysis revealed digital technology self-efficacy, locus of control, and interactions did not significantly predict trust in AI. These findings indicated locus of control and self-efficacy may not significantly influence trust in AI as previously expected, highlighting the need to explore additional psychological or contextual factors affecting user attitudes. By encouraging continued investigation, these findings build a foundation for enhancing user confidence and supporting equitable AI adoption, thereby contributing to positive social change.

Moderating Role of Locus of Control on the Relationship Between Self-Efficacy and

Trust in AI

by

Amelia Gillies

MPhil, Walden University, 2023

PGDipEd, Wesley Institute (Excelsia College), 2010

BA, Monash University, 2006

Dissertation Submitted in Partial Fulfillment

of the Requirements for the Degree of

Doctor of Philosophy

Psychology

Walden University

May 2025

Table of Contents

List of Tables

List of Figures

Chapter 1: Introduction to the Study

The rapid advancement of artificial intelligence (AI) technologies has transformed numerous industries, including healthcare, transportation, and customer service. Among various AI applications, AI chatbots and conversational agents have become increasingly prominent, offering assistance across multiple sectors (Esteva et al., 2019; To et al., 2021). As AI systems take on more complex tasks with the potential to significantly influence decision-making, the issue of trust in AI has emerged as a critical factor for their successful integration and use (Hancock et al., 2011). Without sufficient trust, users may avoid or underutilize AI systems, diminishing potential benefits these technologies can offer. Therefore, understanding psychological factors that influence trust in AI is essential for fostering greater engagement and confidence in AI-driven tools.

Trust in AI is shaped by a variety of psychological and external factors. Two key psychological traits, self-efficacy and locus of control, play important roles in terms of how individuals perceive and engage with technology more broadly (Bandura, 1977; Rotter, 1966). Recent studies have applied them to the context of trust in AI. Self-efficacy refers to individual belief in ability to effectively use technology (Bandura, 1977), while locus of control pertains to individual perceptions of personal control over outcomes (Rotter, 1966). Individuals with high self-efficacy are more likely to trust AI, as they feel confident in their abilities to use these systems (Hancock et al., 2011). Conversely, individuals with a strong internal locus of control might exhibit caution or resistance to fully autonomous AI systems, perceiving them as potential threats to their preference for maintaining direct oversight and decision-making, even though they retain the ability to override AI outcomes (Chiou et al., 2021). This inclination stems from the core aspect of internal locus of control: a belief in one's ability to influence events. While this belief often results in confidence when interacting with new technologies, it can also lead to frustration when perception of control is challenged (Chiou et al., 2021).

Although self-efficacy and locus of control have been studied independently in relation to technology use, the potential for locus of control to moderate the relationship between self-efficacy and trust in AI had not previously been examined. To that end, I examined the extent to which locus of control moderated the relationship between self-efficacy and trust in AI among English-speaking adults between 22 and 55 living in the United States. Through this examination, findings included insights about how individual psychological traits influence user trust in AI, which may inform design of more user-friendly AI systems. These insights may help AI developers and organizations address psychological barriers to AI adoption, potentially leading to more efficient and trustworthy AI applications across various domains.

In this chapter, I provide the background of the problem, problem statement, purpose of the study, research questions and hypotheses, theoretical framework, nature of the study, relevant definitions, assumptions, scope and delimitations, limitations, and the significance of the study.

Background

AI technologies have become deeply embedded in many sectors, including healthcare and transportation (Esteva et al., 2019; Chen et al., 2016). As chatbots and other AI systems take on increasingly complex decision-making roles, user trust in these systems is critical for their effective implementation. However, trust in AI remains a significant challenge, as users grapple with issues such as job displacement and misinformation (Lazer et al., 2018; Piercy & Gist-Mackey, 2021). In this context, locus of control and self-efficacy have emerged as crucial factors influencing trust in AI systems (Hancock et al., 2011). Individuals with a strong internal locus of control and believe they have personal control over their lives may exhibit cautious behavior when interacting with AI systems, particularly in scenarios where these systems are perceived as limiting their direct influence (Chiou et al., 2021). While such individuals often demonstrate confidence in their ability to engage with technology, this may reflect an idiosyncratic response related to fully autonomous decision-making. Individuals with a strong internal locus of control typically approach new technological challenges, including AI, with resilience and problem-solving strategies, maintaining their confidence when mastering technology (Babiker et al., 2024; Wang et al., 2022). This highlights the importance of contextual factors in terms of interpreting locus of control behavior in relation to AI trust.

Self-efficacy, or the belief in one's ability to complete tasks, has been associated with greater trust in technology (Hancock et al., 2011). While trust in AI has been studied across various domains, the extent to which locus of control moderates the relationship between self-efficacy and trust in AI remains underexplored. Locus of control can act as a moderating variable, influencing how individuals' psychological traits interact with their attitudes and behaviors toward technology use. Wang et al. (2022) found internal locus of control enhanced impact of active learning and task crafting on creativity in a work environment involving AI and robotics, indicating individuals with an internal locus of control are more adaptive and proactive. Babiker et al. (2024) demonstrated internal locus of control contributed to positive attitudes and reduced fear toward AI. These findings suggest locus of control can shape interactions between self-efficacy and trust in AI, highlighting the potential relevance of exploring this relationship further.

Trust is central to the successful use of AI technologies, particularly in industries where reliability and accuracy are critical, such as healthcare. Low trust in AI-generated medical advice can lead to patients ignoring or rejecting potentially beneficial interventions (To et al., 2021). In professional settings, lack of trust in AI systems can cause inefficiencies, as individuals may second-guess or override recommendations provided by AI, leading to increased operational costs and reduced productivity (Lee & See, 2004; Guznov et al., 2020). These examples highlight how trust in AI systems is not just a matter of convenience but an essential factor for their effective integration into everyday life.

However, trust in AI can be undermined by various factors, including rapid technological advancements, concerns about job displacement, and the spread of misinformation. As AI systems become more capable, some individuals worry about the risk of being replaced in the workforce, a phenomenon which is referred to as automation anxiety (Piercy & Gist-Mackey, 2021; Vieira, 2020). This anxiety, combined with the fear that AI systems might prioritize efficiency over human welfare, can significantly reduce people's willingness to trust AI. Moreover, the increasing role of AI in content curation and decision-making processes has contributed to the proliferation of misinformation, further eroding public trust in these systems (Lazer et al., 2018). Algorithms that optimize for engagement rather than factual accuracy can inadvertently spread false or misleading information, intensifying public mistrust (Lazer et al., 2018). AI-driven platforms, particularly on social media, are often perceived as credible sources of information despite their limitations in terms of ensuring accuracy (Edwards et al., 2014). This problem is compounded by the phenomenon of AI hallucinations, in which AI systems generate erroneous content that appears credible (Alkaissi & McFarlane, 2023). Furthermore, AI systems can reflect inherent social biases embedded in their training data, such as racial or gender biases, further complicating the issue of trust (Buolamwini & Gebru, 2018; Bolukbasi et al., 2016). Racial biases in image recognition systems and gender biases related to job recommendations have led to concerns regarding AI perpetuating inequality (Barocas & Selbst, 2016). Such biases not only undermine reliability of AI systems but also contribute to a broader social mistrust of AI technologies.

In addition to societal concerns, locus of control and self-efficacy play a pivotal role in terms of shaping user trust in AI. Individuals with a strong internal locus of control are often more skeptical of external systems like AI (Chiou et al., 2021). They may perceive AI as a force that diminishes their personal agency, making them less likely to place trust in its capabilities (Riedl, 2022). Conversely, those with higher levels of self-efficacy tend to exhibit more trust in technological systems, believing they can effectively interact with and benefit from these tools (Hancock et al., 2011). Despite this understanding, the potential for locus of control to moderate the relationship between self-efficacy and trust in AI remains largely unexplored in existing research. This study was designed to fill that gap by providing critical insights regarding factors that influence user trust in AI systems, with the potential to inform the design and implementation of AI technologies in ways that enhance user trust and adoption.

Problem Statement

With the increasing prevalence of AI technologies, particularly in the form of conversational platforms, understanding how users develop trust in these systems has become critical (Hancock et al., 2011; To et al., 2021). As AI systems become more integrated into sectors such as healthcare and customer service, users' trust in these technologies directly impacts their willingness to rely on AI-generated information or recommendations (Lee & See, 2004; Guznov et al., 2020). However, despite the growing societal reliance on AI, psychological mechanisms that govern trust in these systems remain underexplored.

Locus of control plays an important role in shaping attitudes toward technology (Rotter, 1966; Scherer et al., 2019). Individuals with a strong internal locus of control tend to be more skeptical of systems they perceive as external forces, such as AI, which may explain why they are less likely to place trust in these technologies (Chiou et al., 2021; Riedl, 2022). Conversely, self-efficacy, or the belief in one's ability to perform tasks, has been associated with increased trust in technology, as individuals who feel competent are more confident in interacting with these systems (Bandura, 1977; Hsu et al., 2007).

While studies have examined the influence of locus of control and self-efficacy independently, the extent to which locus of control moderates the relationship between self-efficacy and trust in AI has not yet been investigated. This research is particularly timely given the widespread concerns surrounding AI technologies, including issues of misinformation and inherent societal biases (Lazer et al., 2018; Buolamwini & Gebru, 2018). Understanding how psychological traits influence trust is crucial, as AI systems increasingly impact decision-making in critical domains such as healthcare, where trust is essential for the effective use of AI-driven tools (Esteva et al., 2019). Ultimately, this study aimed to contribute to a more trustworthy, user-friendly AI environment by identifying psychological factors with the potential to inform user education and system design strategies.

Purpose of the Study

The purpose of this quantitative, cross-sectional, correlational study was to determine the extent to which locus of control (moderator/independent variable [IV]) moderates the relationship between self-efficacy (independent variable [IV]) and trust in AI (dependent variable [DV]).

Research Questions and Hypotheses

The following research questions and hypotheses guided this study:

RQ1: Does locus of control predict trust in AI?

$H_0$1: Locus of control does not predict trust in AI.

H_a1: Locus of control does predict trust in AI.

RQ2: Does self-efficacy predict trust in AI?

$H_0$2: Self-efficacy does not predict trust in AI.

H_a2: Self-efficacy does predict trust in AI.

RQ3: Does locus of control moderate the relationship between self-efficacy and trust in AI?

$H_0$3: Locus of control does not moderate the relationship between self-efficacy and trust in AI.

H_a3: Locus of control does moderate the relationship between self-efficacy and trust in AI.

Theoretical Framework

This study was informed by two social psychological theories: Bandura's self-efficacy theory and Rotter's locus of control theory. Bandura (1977) stated individuals' beliefs in their capabilities significantly influence their actions, motivations, and emotional wellbeing. Self-efficacy plays a critical role in terms of how individuals engage with emerging technologies like AI. Hancock et al. (2011) found higher levels of self-efficacy predict greater trust in and use of technology, demonstrating how personal belief in one's abilities directly impacts engagement with AI. Additionally, Marakas et al. (1998) stated computer self-efficacy was important in terms of understanding technology adoption and behaviors in digital environments.

Individuals with higher self-efficacy not only demonstrate greater trust in AI systems, but also tend to engage in more complex interactions with these technologies. Users with high self-efficacy are more likely to explore advanced features of AI systems, take risks when adopting new AI tools, and demonstrate persistence with solving problems when AI-generated outcomes are not immediately clear (Hsu et al., 2007). Conversely, individuals with low self-efficacy may avoid interacting with AI or limit their use to simple tasks, fearing they lack skills to manage more sophisticated AI applications. This avoidance behavior underscores the role of self-efficacy in terms of shaping not just trust, but also depth of interactions with AI technologies.

Rotter's locus of control theory differentiates between individuals with an internal locus of control who believe they control their own outcomes and those with an external locus of control who believe their lives are controlled by external factors. This theory is vital to understanding how individuals approach trust in AI. Individuals with a strong internal locus of control often feel confident in terms of their abilities to manage external influences, including AI systems (Chiou et al., 2021; Riedl, 2022). Chiou et al. (2021), suggested individuals with an internal locus of control may exhibit caution or skepticism toward AI systems when these technologies are perceived as potentially limiting personal decision-making autonomy. contrasts with the broader body of research indicating individuals with an internal locus of control typically approach new challenges, including AI, with resilience and confidence in terms of maintaining their decision-making autonomy (Babiker et al., 2024; Wang et al., 2022). Conversely, those with higher levels of self-efficacy are more likely to trust AI systems, as their belief in their own abilities makes them confident when interacting with and managing AI technologies (Bandura, 2006; Hsu et al., 2007).

The degree to which individuals exhibit skepticism or acceptance of AI systems is often shaped by their locus of control. Individuals with a strong internal locus of control may question AI-driven decisions, as they prefer to retain control over critical decisions rather than delegate to automated systems (Chiou et al., 2021). This may appear to contradict conventional interpretations of internal locus of control, which suggest confidence in one's ability to handle external systems or influences. Chiou et al. (2021) suggested individuals with an internal locus of control might view AI systems as external entities that challenge their autonomy and personal decision-making authority. Individuals with an internal locus of control generally maintain confidence and adaptability when engaging with AI, viewing it as a tool they can master rather than an external force that undermines their autonomy (Babiker et al., 2024; Wang et al., 2022).

By contrast, those with an external locus of control who doubt their control over outcomes may defer to perceived greater competence and accuracy of AI technologies when handling decision-making (Riedl, 2022). Such individuals are more likely to integrate AI use while maintaining a proactive stance toward decision-making (Wang et al., 2022).

Self-efficacy and locus of control theories were suitable to explain the moderating role of locus of control on the relationship between self-efficacy and trust in AI, which is critical to understanding variations in terms of user behaviors. Individuals with high self-efficacy and an internal locus of control are likely to feel confident in their abilities but also skeptical about relying on AI for important decisions. This group may cross-check AI outputs due to their own judgments or make final decisions independently. Conversely, individuals with high self-efficacy and an external locus of control may demonstrate high confidence in their skills but be more willing to rely on AI as they perceive it as a competent external resource. These behavioral differences highlight complex dynamics involving psychological traits and trust in AI, suggesting locus of control can influence whether high self-efficacy leads to active engagement or cautious interaction with AI technologies. I explain the self-efficacy and locus of control theories in more detail in Chapter 2.

To test the moderating effect of locus of control on the relationship between self-efficacy and trust in AI, I used interaction effects via a multiple regression framework. By measuring self-efficacy and locus of control independently and assessing their relationships with trust in AI, I determined whether locus of control strengthened or weakened associations between self-efficacy and trust. Specifically, I examined whether individuals with a high internal locus of control were less likely to trust AI even when their self-efficacy was high as compared to individuals with an external locus of control who might exhibit more trust in AI despite varying levels of self-efficacy. I aimed to provide a nuanced understanding of how psychological traits interact to shape user behaviors in relation to AI technologies.

Nature of the Study

I employed a quantitative cross-sectional nonexperimental, correlational survey design. This design was appropriate for examining relationships between self-efficacy, locus of control, and trust in AI. A correlational design does not permit causality to be inferred. The quantitative approach was used to measure and analyze these variables and their interactions, making it suitable for assessing the topic. A cross-sectional design was ideal for gathering data at a single point in time from a large sample, ensuring efficiency in terms of examining these relationships without manipulating any variables.

The target population was English-speaking adults between 22 and 55 living in the United States. Participants were recruited via Amazon's Mechanical Turk (MTurk). Data were analyzed using Statistical Package for Social Sciences (SPSS) Version 29.0. A multiple regression analysis with moderation was conducted to address the topic.

Definitions

Artificial Intelligence (AI): Systems and machines that simulate human intelligence in order to perform tasks, learn from data, and improve over time (Russell & Norvig, 2010).

AI Hallucinations: Instances where AI systems, particularly those that are driven by machine learning, generate false or misleading information that appears credible. These hallucinations can result from errors in data processing or due to the model producing responses based on incomplete or inaccurate information (Alkaissi & McFarlane, 2023). Such errors are particularly relevant in AI systems that are used for decision-making or customer interactions.

AI Interaction Frequency (Regular AI Interaction): Regular AI interaction is a measure of as use of AI-driven tools or systems (e.g., smart assistants, AI-powered customer service platforms) at least several times per week. This ensured participants in the study had relevant experience with AI technologies.

Automation Anxiety: Fear or concern that AI and automation technologies will replace human jobs, leading to job displacement and/or reduced human roles in certain sectors. This anxiety can reduce individuals' willingness to trust or adopt AI technologies (Piercy & Gist-Mackey, 2021).

Locus of Control: The extent to which individuals believe they can control events that affect their lives. Individuals with an internal locus of control believe they can influence outcomes through their own actions, while those with an external locus of control attribute outcomes to external factors that are beyond their control (Rotter, 1966).

Self-Efficacy: Individual belief in their ability to execute tasks and achieve specific outcomes (Bandura, 1977).

Technology Self-Efficacy: Individual confidence in their ability to use and adapt to new technologies (Marakas et al., 1998).

Trust in AI: Individual willingness to rely on AI systems to perform tasks that have meaningful consequences.

Assumptions

There were several assumptions associated with this study. I assumed participants provided honest and accurate responses to survey questions. As I relied on self-reported data, it could not be directly verified if participants were truthful. To mitigate potential dishonesty, participants were assured their responses were anonymous, which was expected to encourage honesty. Additionally, I assumed participants engaged thoughtfully with survey questions and dedicated sufficient time to providing considered responses. This was crucial to ensure data accurately reflected their experiences and perceptions. I also assumed locus of control and self-efficacy theories were suitable to inform the study.

Scope and Delimitations

This study involved determining the extent to which locus of control moderates the relationship between self-efficacy and trust in AI. I focused on understanding user engagement with AI technologies. Both self-efficacy and locus of control were independently related to engagement with and trust in AI technology. However, the potential for locus of control to moderate the relationship between self-efficacy and trust in AI had not previously been examined empirically. While other factors might influence trust in AI, such as demographic variables or prior experience with technology, I limited the scope of this study to psychological traits involving comfort with and/or resistance to AI technology use. Participants were English-speaking adults between 22 and 55 residing in the United States. To gain insights regarding their experiences with AI, a demographic questionnaire was used to collect information on types of AI technologies (smart assistants, AI-powered customer service platforms, smartphone applications) and frequency of use.

Limitations

Convenience sampling through MTurk was used to recruit voluntary participants, which might have resulted in a nonrepresentative sample in terms of those who did and did not volunteer. The volunteer-based nature of sampling potentially limited representativeness of the sample and generalizability of findings.

Another limitation was reliance on self-reported data, which could have been influenced by social desirability bias; participants might have responded in ways they believed were more socially acceptable rather than being entirely truthful. Reminding participants that all responses were collected anonymously was intended to mitigate this by reducing pressure to provide socially desirable answers.

Significance

This study contributed to the field of social psychology by exploring psychological mechanisms that influence trust in AI. Specifically, I addressed a gap in literature by examining whether locus of control moderated the relationship between self-efficacy and trust in AI. Although previous research had examined direct effects, the moderating role of locus of control had not previously been empirically tested. Findings included initial insights regarding the relationship between these psychological traits and trust in AI, suggesting directions for future research.

In terms of practical implications, this research includes preliminary insights that could be valuable for AI developers, technology companies, and organizations. While I did not identify significant relationships, exploration of self-efficacy and locus of control in relation to trust in AI could guide further investigation into designing user-friendly AI systems. These preliminary findings suggest areas for future research, potentially contributing to improved training programs and user interfaces that are aimed at boosting user confidence. Organizations may use insights from this exploratory work to inform further research or pilot studies for understanding psychological barriers to AI integration in the workplace. Encouraging trust in AI remains an important goal, as increased trust may enhance effectiveness of AI-driven tools, potentially improving outcomes in healthcare and customer service. While findings from this study were not significant, exploration of psychological factors influencing trust in AI highlights leads to further research that could ultimately benefit individuals and society. Addressing trust in AI remains essential for positive social change, particularly as AI becomes more prevalent in healthcare, education, and customer service. By exploring psychological traits related to trust in AI, I identified potential avenues for future research that could eventually inform policies and practices promoting broader acceptance and responsible AI use.

Summary

Rapid development of AI technologies has significantly impacted various industries, such as healthcare, transportation, and customer service, making trust in AI systems an essential factor for successful integration and use. This study involved addressing self-efficacy and locus of control, which may affect how individuals engage with and trust these technologies. I examined whether locus of control moderated the relationship between self-efficacy and trust in AI among English-speaking adults between 22 and 55 living in the United States. By using a quantitative cross-sectional correlational design, I aimed to address this topic with the broader goal of informing design and implementation of user-friendly AI systems.

In Chapter 2, I address the theoretical frameworks and key psychological concepts. I provide an exhaustive review of literature related to self-efficacy, locus of control, and trust in AI.

Chapter 2: Literature Review

With the rise in popularity of A technology, particularly organization platforms with human-like conversational abilities, societal reliance on such technology has grown, elevating the significance of misinformation and AI hallucinations as social issues (Bickmore & Picard, 2005; Ciechanowski et al., 2019). In navigating the complex landscape of AI and related challenges, the presence of misinformation and ungrounded, erroneous content raises significant concerns. These AI platforms which are often perceived as credible can generate and spread information that may not be based on factual data, intensifying the social dilemma of using online misinformation. This problem is further compounded by difficulties many individuals experience when discerning between human and AI-generated text, amplifying the spread and acceptance of misinformation (Kreps et al., 2022). Consequently, understanding these dynamics as part of the broader social problem of AI-driven misinformation is imperative.

AI systems are not immune to inherent social biases. These biases include but are not limited to racial, ethnic, gender, and socioeconomic biases. For instance, AI systems can exhibit racial biases by favoring majority racial groups via image recognition or perpetuating stereotypes that reflect biases in their training data (Buolamwini & Gebru, 2018). Gender bias is another significant issue, with AI systems often associating certain professions or roles with a specific gender, mirroring biases in training data (Bolukbasi et al., 2016). These biases can influence the type and tone of information, further complicating issues involving misinformation (Barocas & Selbst, 2016; Howard & Borenstein, 2018).

The complex interplay of social biases and misinformation is particularly significant in the context of social media platforms. The sheer volume of information, its rapid dissemination, and reliance on AI systems for content moderation make social media particularly prone to misinformation and bias propagation. Algorithms may unwittingly lead to echo chambers, reinforcing users' existing beliefs and biases and contributing to the polarization of social discourse (Bakshy et al., 2015). Moreover, the potential role of AI in media manipulation in terms of creation and dissemination of deepfakes and sophisticated digital forgeries underscores the urgent need to address this social problem (Chesney & Citron, 2018; Groh et al., 2021). Therefore, understanding how user trust in AI systems is formed and maintained is essential for navigating this complex and pressing social problem.

Relevance of the Problem

The ever-increasing integration of AI into daily lives presents both promises and pitfalls. While AI platforms, especially those with humanlike conversational abilities, offer the potential for societal advancement, they also raise considerable concerns. These concerns involve misinformation and the perpetuation of biases (Alkaissi & McFarlane, 2023; Bickmore & Picard, 2005; Ciechanowski et al., 2019; Salvagno et al., 2023). As a result, public trust in AI becomes a critical factor that demands comprehensive examination.

Lack of trust in AI leads to concrete barriers that undermine its effective integration across various sectors. In healthcare, distrust can cause patients to ignore AI-driven medical advice, while in collaborative settings, it can lead to inefficient human-AI partnerships and increased operational costs (Guznov et al., 2020; Lee & See, 2004; To et al., 2021). This distrust is further fueled by automation anxiety, contributing to job insecurity and reduced willingness to interact with AI in professional settings (Piercy & Gist-Mackey, 2021; Vieira, 2020). Misinformation perpetuated by AI may also corrode trust, affecting healthcare and news dissemination (Chou et al., 2018; Lazer et al., 2018). Social biases and misinformation particularly on social media platforms further complicate trust in AI (Bakshy et al., 2015; Chesney & Citron, 2019; Groh et al., 2021). Rapid pace of AI advancements exacerbates the digital divide, eroding trust among those who feel overwhelmed or marginalized by these technological changes (Van Dijk, 2017).

A rich body of literature highlights the influential role of locus of control and self-efficacy on shaping attitudes towards technology and fostering trust. Rotter's concept of locus of control, or the degree to which individuals perceive control over events affecting them, is a significant predictor of trust in technology. Scherer et al. (2019) stated those with a strong internal locus of control, tend to have greater trust in technology. Bandura's self-efficacy also plays a critical role in terms of fostering trust in technology. Individuals with high self-efficacy levels are likely to exhibit more confidence during interactions with technology, leading to increased trust (Bandura, 2006; Hsu et al., 2007).

The relationship between locus of control, self-efficacy, and trust in AI is complex and shaped not only by AI proficiency and accuracy of information but also human users' perceptions and psychological predispositions. By investigating the interplay of these variables, I aimed to provide insights that could potentially inform user education strategies, contributing to efforts to foster more trustworthy AI environments.

This research will lead to understanding psychological traits that are relevant to trust in AI, potentially guiding future efforts for mitigating harms of AI-generated misinformation and biases.

The purpose of this quantitative study was to examine the degree to which locus of control (IV) moderated the relationship between self-efficacy (IV) and trust in AI (DV). Given the impacts of AI-driven misinformation and inherent societal biases, understanding how self-efficacy and locus of control influence trust in AI is paramount.

Chapter 2 includes information about the literature search strategy followed by a detailed description of the theoretical frameworks. An exhaustive review of literature related to locus of control, self-efficacy, and trust in AI is presented, followed by a summary and conclusion.

Literature Search Strategy

I used the following databases for this study: EBSCOHost, Science Direct, Thoreau Multi-Database Search, APA PsycInfo, and APA PsycArticles. Keywords were: *Artificial intelligence*, *AI chatbots*, *AI hallucinations*, *human-computer interaction*, *AI trust*, *self-efficacy and AI*, *locus of control and AI*, *AI bias*, *AI and social inequality*, *uncanny valley effect*, *human-chatbot interaction*, *AI in clinical interviews*, *virtual humans*, *machine behavior*, *user expectations of AI*, *AI and media manipulation*, *AI in scientific writing*, *misinformation and AI*, *AI and social media*, *AI and echo chambers*, *AI and deepfakes*, *image recognition and bias*, *racial bias in AI*, *gender bias in AI*, *socioeconomic bias in AI*, *AI content moderation*, *individual traits and AI trust*, *robotics and trust*, *trust in automated systems*, *self-efficacy and technology*, *psychological factors* *and AI trust*, *Bandura and AI*, *Rotter and AI*, *technology adoption*, *trust mechanisms in AI*, *computer self-efficacy*, *locus of control in technology*, *human-AI interface*, *AI ethics*, *AI in healthcare*, *AI in mental health*, *social implications of AI*, *trust in technology*, and *user engagement in AI*.

I focused on sources that were published between 2020 and 2025, excepting seminal works and information about theoretical frameworks with enduring applicability. I facilitated the gathering of a comprehensive and multidisciplinary body of research with diverse viewpoints and methodologies to richly inform this study.

Theoretical Foundation

Self-Efficacy Theory

Origin and Source

Bandura introduced the concept of self-efficacy in 1977 as an integral component of his broader social cognitive theory. While the psychological theories at that time offered significant perspectives on learning and motivation, many did not adequately consider the role of individual agency. Bandura's social cognitive theory bridged this gap by introducing a model of causality that interwove the individual, their behaviors, and their environment in a dynamic referred to as *reciprocal causation* (Bandura, 1986). Self-efficacy specifically served to articulate how people's beliefs in their abilities could profoundly affect their actions and outcomes within this complex social system.

Anchored in the broader landscape of cognitive psychology, self-efficacy theory represents a paradigmatic shift that began to gather momentum in the mid-20th century. Unlike behaviorism, which was then the dominant psychological framework, cognitive psychology emphasized internal processes such as thought, memory, and decision-making (Neisser, 1967). This focus allowed scholars to move beyond examining only the stimulus-response mechanisms and explore the internal cognitive structures that mediate between stimulus and response (Miller et al., 1960). Bandura's conceptualization of self-efficacy perfectly aligns with this shift by emphasizing the role of cognitive processes in understanding human behavior, such as how beliefs about one's capabilities can influence actions and emotional states (Bandura, 1997).

Major Propositions and Hypotheses

Self-efficacy is not a peripheral factor within Bandura's seminal works, but a central construct that significantly influences various dimensions of psychological functioning and human behavior (Bandura, 1997). A key proposition of the theory is the emphasis on how self-efficacy shapes the way in which individuals choose to participate in different activities. For example, individuals with higher self-efficacy in technology usage are more likely to trust AI systems, engaging more confidently with AI-driven platforms.

Bandura asserts that people are more likely to engage in tasks where they perceive a higher level of competence driven by their belief in their ability to achieve the desired outcomes (Bandura, 1982). Given that tech proficiency is, and will continue to be, required for employment in most areas, trust in and comfort with AI technology will serve individuals very well, both professionally and personally. In essence, the higher the level of perceived self-efficacy, the more likely an individual is to undertake more challenging tasks. This is consistent with the theory's cognitive underpinnings, where belief systems guide action. Within this context, belief systems refer to internal frameworks individuals use to evaluate their skills and abilities. These frameworks are not merely passive but serve as active determinants that influence decision-making and behavior, functioning as lenses through which individuals assess the situation, their own competencies, and the likely outcomes of different courses of action (Bandura, 1982). This suggests that individuals with higher tech self-efficacy in technology might more confidently assess their ability to use AI, leading to greater trust in AI systems and the proficiency required to succeed in whatever the endeavor.

According to empirical studies, individuals' level of effort and persistence in tasks correlates strongly with their self-efficacy beliefs (Schunk, 1991). The reason lies in the self-regulatory function of self-efficacy; people with high self-efficacy are more likely to set challenging goals for themselves and invest the necessary effort to achieve them. This is substantiated by Bandura's assertion that elevated levels of self-efficacy inspire greater effort and resilience when confronted with obstacles (Bandura, 1997). For example, individuals with high self-efficacy in technological skills are more likely to actively engage with and trust AI systems, persisting through challenges in understanding and using these technologies effectively, persistence that will serve them well at work and elsewhere. In practical terms, individuals endowed with high self-efficacy are less prone to discouragement in the face of setbacks and more likely to sustain effort over prolonged periods. This observation aligns well with the principles of humanistic psychology, which values personal growth and resilience. In this context, self-efficacy serves as a psychological mechanism that fuels an individual's motivation, perseverance, and adaptability (Bandura, 1982).

Beyond its influence on task choices and effort, self-efficacy significantly impacts emotional well-being (Bandura, 1997). Research suggests that heightened levels of self-efficacy serve as protective factors against stress and depression (Jerusalem & Schwarzer, 1992). Specifically, individuals with robust self-efficacy beliefs perceive stressful situations as challenges rather than threats, thereby reducing anxiety and enabling better coping strategies. Individuals with higher self-efficacy are observed to employ more effective problem-solving techniques when confronted with challenging situations, further enhancing their coping abilities (Bandura, 1997). Self-efficacy theory thus adds a nuanced layer to our understanding of emotional resilience, emphasizing the role of belief systems in shaping cognitive and emotional responses to stress (Bandura, 1982).

Bandura also elaborated on the mechanisms through which self-efficacy beliefs are formed and nourished. He identified four primary sources: *mastery experiences* (i.e., individuals gain a sense of accomplishment), *vicarious experiences* (i.e., observing others succeed boosts one's own sense of efficacy), *verbal persuasion* (i.e., encouragement and constructive feedback foster a greater belief in one's capabilities), and *physiological states* (i.e., emotional states and physical reactions can either enhance or diminish self-efficacy; Bandura, 1986). For example, these mechanisms imply that individuals might develop trust through successful personal interactions with AI (mastery experiences), by observing others effectively use AI (vicarious experiences), receiving positive reinforcement about their ability to use AI (verbal persuasion), and from their emotional responses to AI interactions (physiological states).

Each of these sources not only enriches our understanding of how self-efficacy is formed but also emphasizes the theory's multidimensional nature, underscoring its dynamic nature, shaped and reshaped by a complex interplay of both personal achievements and external influences (Bandura, 1986).The propositions and hypotheses within self-efficacy theory underscore its comprehensive scope, offering robust explanations for individual choice in activities, effort allocation, emotional resilience, and the complex mechanisms of belief formation.

Previous Applications of the Self-Efficacy Theory in Similar Studies

Pajares (1996) aimed to explore the predictive and mediational roles of self-efficacy beliefs in mathematical problem-solving among middle school gifted students mainstreamed with regular education students in algebra classes. The authors hypothesized that self-efficacy beliefs of gifted students would make an independent contribution to problem-solving performance. Pajares employed path analysis on a sample of 297 eighth-grade students, including both regular education (n = 231) and gifted students (n = 66), to test this hypothesis. The participants were asked to complete a descriptive questionnaire (Pajares, 1996), the self-efficacy for self-regulated learning measures scale (Zimmerman et al., 1992), and a math based self-efficacy instrument (Pajares & Miller, 1994). Findings indicated that gifted students reported higher math self-efficacy and self-efficacy for self-regulated learning and lower math anxiety compared to regular education students. Pajares concluded that self-efficacy of gifted students significantly predicted their problem-solving abilities, independent of factors like math anxiety, cognitive ability, mathematics GPA, self-efficacy for self-regulated learning, and gender.

The results supported the hypothesized role of self-efficacy in Bandura's social cognitive theory, indicating that self-efficacy beliefs are strong predictors of capability in academic tasks, particularly in mathematical problem-solving. These findings are relevant to the proposed study, as self-efficacy theory has the potential to explain individuals' trust in and ability to use AI effectively. Pajares' (1996) exploration of self-efficacy in academic problem-solving establishes a foundational framework that Mlekus et al. (2020) build upon. While Pajares illuminates the role of self-efficacy in structured academic tasks, Mlekus et al. extend this concept to the realm of technology acceptance. This progression reflects an evolution of self-efficacy from a predictor of academic success to a crucial factor in user interactions with new technologies, like R software.

Mlekus et al. (2020) extended the Technology Acceptance Model originally developed by Davis (1989), which focuses on perceived usefulness and ease of use as determinants of technology acceptance. They expanded the Technology Acceptance Model by incorporating user experience characteristics, such as efficiency, perspicuity, dependability, stimulation, and novelty, to provide a more comprehensive understanding of technology acceptance. The authors hypothesized that specific user-experience characteristics, such as efficiency, perspicuity, dependability, stimulation, and novelty, would significantly predict technology acceptance, including perceived usefulness and ease of use. The study involved 281 statistics students, mostly female (*M* age = 23.29

years), who had recently learned to use R software, an open-source programming language and software environment widely used for statistical computing and graphics. The participants completed the User Experience Questionnaire (Laugwitz et al., 2008) and scales from TAM3 (Venkatesh & Davis, 2000), which include measures for perceived usefulness, ease of use, and behavioral intention.

The results indicated that user-experience characteristics accounted for significant variance in technology acceptance measures. Output quality (i.e., the degree to which users believe that the system performs their job tasks well) significantly predicted perceived usefulness of the technology, while perspicuity and dependability were significant predictors of perceived ease of use. Additionally, novelty (but not stimulation) was a significant predictor of behavioral intention to use the technology. These findings support the authors' hypothesis and emphasize the importance of user-experience characteristics in technology acceptance. These findings underscore how user confidence and competence in using a technology (in this case, the R software) are influenced by their experiences with its technological features, enhancing users' beliefs in their capabilities to accept and effectively use, and continue to use, the technology. The findings of Mlekus et al. (2020) about self-efficacy in technology usage set the stage for Lestari et al. (2022), who explore self-efficacy in a professional environment impacted by AI and robotics. Here, self-efficacy transcends the boundaries of individual user experience to influence organizational behavior and job performance in the hotel industry. This transition highlights the broadening scope of self-efficacy theory from personal competence to organizational effectiveness in technology-rich settings.

Lestari et al. (2022) conducted a study to investigate the influence of AI, robotics, and service automation on job performance in the hotel industry; its purpose was to examine how employees' self-efficacy, attitudes toward technology adoption, and relationship quality with supervisors' impact job performance amid AI advancements. The authors hypothesized that self-efficacy would positively influence job performance and attitudes towards technology adoption. The study involved 171 employees from eight large premium hotels in Jakarta, Indonesia. These employees, who regularly interacted with technology equipment or systems, completed a questionnaire that included the self-efficacy scale developed by Schyns and Von Collani (2002), performance measures derived from Ang et al. (2003), and the attitude towards technology adoption scale introduced by Williams et al. (2017). The research utilized a questionnaire survey as its primary method, which included items to measure self-efficacy using the scale developed by Schyns and Von Collani (2002), as well as measures for employee performance and attitudes toward technology. The questionnaire utilized a Likert-type scale for responses, where self-efficacy was measured using Schyns and Von Collani's (2002) six-item instrument, performance was gauged using six items from Ang et al. (2003), and attitudes towards technology were measured using Williams et al.'s (2017) scale.

The study's results indicated a positive association between employees' self-efficacy and their job performance. Additionally, employees' attitudes toward smart technology adoption positively influenced their job performance. Additionally, the study found that employees' self-efficacy positively influenced their attitudes toward smart technology adoption and attitudes toward smart technology adoption were found to positively influence their job performance, thus supporting all three hypotheses. The results of Lestari et al. emphasize the significant role of self-efficacy in employees' performance and attitudes towards technology adoption, particularly in environments with advanced technologies like AI and robotics. These findings are relevant to the proposed study, highlighting the importance of self-efficacy in professional settings influenced by AI advancements. Lestari et al. (2022) establish the relevance of self-efficacy in professional settings influenced by AI advancements, paving the way for Al-Gasawneh et al. (2023) to delve into the more intricate domain of IoT and SEO. Al-Gasawneh et al.'s focus on IoT's impact and user behavior in digital marketing contexts showcases the expanding applicability of self-efficacy theory. It demonstrates its critical role in navigating the complex interplay between advanced technological systems and user adaptation, aligning well with the evolving nature of modern digital environments.

Finally, Al-Gasawneh et al. (2023) examined the impact of the Internet of Things (IoT), a network of interconnected devices that collect and exchange data, on Search Engine Optimization (SEO; i.e., the practice of increasing the visibility of websites in search engine results). The study specifically focused on the roles of user behavior and self-efficacy in this relationship. The authors hypothesized that the direct influence of IoT on SEO and user behavior and the mediating role of user behavior in the IoT-SEO relationship. This mediation effect was predicted to mean that changes in user behavior, influenced by the introduction and use of IoT technologies, would in turn affect SEO outcomes. The study posits that IoT advancements would influence how users interact with search engines, which would then impact SEO effectiveness. Additionally, it examined how users' self-efficacy might moderate the relationship between IoT and user behavior, by examining how users' self-efficacy could influence the relationship between IoT and user behavior. It was hypothesized that users with higher self-efficacy would be more adept at adapting to and utilizing IoT technologies, thereby positively influencing their behavior in ways that could enhance SEO outcomes. In other words, the moderation effect suggests that the strength of the relationship between IoT and user behavior varies depending on the level of users' self-efficacy. The research employed a survey method, targeting IT and digital marketing professionals working in Jordanian telecommunications companies. A total of 160 professionals were approached via email, and 131 valid and complete responses were ultimately analyzed. The survey measured variables such as performance expectations, social impact, cost, security (to capture IoT), user behavior, and users' self-efficacy. All items were rated on a 5-point Likert scale and underwent field and academic pilot testing for clarity, applicability, and validity.

The study's findings revealed significant relationships between IoT, user behavior, and SEO. Specifically, IoT was found to have a direct positive impact on SEO and user behavior. The "direct impact" of IoT on SEO was independent of its impact on user behavior, meaning IoT directly affects both SEO and user behavior separately. The study suggests that IoT technologies not only influence how users behave but also directly enhance SEO effectiveness by enabling more efficient and targeted search engine optimizations. Moreover, user behavior significantly mediated the relationship between IoT and SEO, indicating that changes in user behavior due to IoT, in turn, affected SEO outcomes. Results also revealed that users' self-efficacy moderated the association between IoT and user behavior, highlighting the importance of individuals' confidence in their abilities in the context of technology use and adaptation. The moderation effect observed in the study indicated that higher self-efficacy among users strengthens the relationship between IoT and user behavior. In essence, when users have greater confidence in their ability to use IoT technologies, there is a more pronounced positive impact of IoT on their behavior. This suggests that user self-efficacy is a critical factor in maximizing the benefits of IoT in influencing user behavior and, by extension, SEO outcomes. These findings are relevant to the current research focusing on trust in AI systems, as they demonstrate the importance of self-efficacy in navigating and adapting to emerging technologies like IoT, which can have significant implications for user behavior and perceptions, particularly in digital marketing and SEO strategies.

Rationale for Use of the Self-Efficacy Theory

The rationale for selecting self-efficacy theory as the theoretical foundation for this study was rooted in its relevance to understanding individual interactions with AI, particularly in the context of the increasing complexity and prevalence of misinformation. Self-efficacy theory, as articulated by Bandura (1977, 1986), offers a robust framework for examining how individuals' beliefs in their abilities influence their behaviors and attitudes, especially in technologically driven environments.

The theory's emphasis on the role of individual agency in shaping interactions with the environment was particularly pertinent to the study's focus on AI, where users are constantly interacting with advanced, often opaque systems; users' beliefs in their ability to effectively navigate these systems are likely to significantly influence their trust in AI. This aligns with Bandura's (1986) assertion that self-efficacy beliefs shape how individuals approach tasks, challenges, and decision-making processes. Additionally, the multifaceted nature of self-efficacy theory, which encompasses cognitive, motivational, affective, and selection processes (Bandura, 1986, 1997), provided a comprehensive lens through which to examine the study's focal issues. Given that trust in AI involves cognitive evaluations of the technology's capabilities, motivational aspects of choosing to rely on AI, and affective responses to AI-generated information, self-efficacy theory aptly encapsulates these dimensions.

How the Self-Efficacy Theory Relates to The Present Study

Self-efficacy theory (Bandura 1977, 1986) provides a comprehensive framework for understanding how individuals' beliefs in their capabilities influence their interactions with AI, an aspect that is crucial in an era dominated by digital technologies and pervasive misinformation. The specific research questions of this study were designed to both apply and extend Bandura's concepts by examining how self-beliefs (e.g., locus of control) shape behavior. The relationship between self-efficacy and trust in AI resonates with Bandura's (1997) findings on the pivotal role of self-efficacy in shaping individual interactions with various aspects of their environment, including technology. This aspect of the study underscores the relevance of self-efficacy in the digital age, as affirmed by numerous studies linking self-efficacy with effective engagement with technology (Hsu et al., 2007; Lucas et al., 2014).

Literature Review Related to Key Variables

Trust in AI

In the context of AI, trust extends beyond mere reliability; it encapsulates the belief in AI's capability to act as credible social actors within a vast array of fields, from healthcare and news dissemination to political discourse and software development. The pervasive integration of AI into these sectors, particularly in areas where trust is paramount, drives the urgent need to address challenges posed by misinformation and AI hallucinations (i.e., ungrounded or erroneous content generated by AI systems) that exacerbate the complexity of fostering trust in these technologies. The articles underpinning this section have been curated to illuminate the diverse aspects of trust in AI across various domains. It aims to showcase the intricate relationship between AI technologies and societal trust, emphasizing AI's critical role as social actors embedded within the fabric of daily life. By navigating through the theoretical foundations of trust in technology, empirical evidence on trust in automation and AI, and the nuanced impacts of misinformation, societal biases, and psychological traits on trust, this section endeavors to construct a comprehensive understanding of the dynamics at play. It underscores the imperative to discern the mechanisms through which AI technologies can be perceived as trustworthy amidst the challenges of misinformation and inherent societal biases. These articles collectively contribute to a nuanced discourse on trust in AI, highlighting its significance in the digital age and laying the groundwork for investigating the moderating role of locus of control in shaping trust in AI technologies.

Chien et al. (2016) offered a foundational study that underscores how cultural and personality differences critically impact trust in automation technologies. Their research across three different countries highlights the significant role of cultural traits such as *uncertainty avoidance* and *individualism*, alongside personality traits like *agreeableness* and *conscientiousness*, in influencing trust levels. This insight broadens the discourse from AI's functional reliability to include cultural and psychological dimensions that shape user interactions with technology.

Chien et al. (2016) investigated the influence of cultural characteristics and personality traits on trust in automation across three distinct populations: the United States, Taiwan, and Turkey by examining the interplay between Hofstede's cultural dimensions, the Big Five personality traits, and their collective impact on trust attitudes toward automation. The authors employed several measurement tools to achieve this: the Culture Trust Instrument (CTI) (Chien et al., 2016), Hofstede’s Cultural Dimensions (Hofstede, 1991) to evaluate cultural variances, and the Big Five Personality Traits scale (John & Srivastava, 1999). The study sampled 360 participants, with 120 individuals from each country. The demographic distribution included various age groups and both genders. The results revealed significant differences in trust levels, with the US participants displaying the highest trust in automation, followed by the Taiwanese participants, and the Turkish participants showing the least. This variance was significantly influenced by cultural dimensions such as uncertainty avoidance and individualism, which correlated positively with trust levels in the US. Additionally, personality traits like agreeableness and conscientiousness were positively related to trust across all groups. These findings deepen our understanding of trust in AI by emphasizing how cultural and personality differences significantly impact user perceptions and acceptance, underscoring the importance of designing AI systems that are adaptable to various cultural contexts and personality types to foster trust among a diverse range of global users.

Building on these insights, Oksanen et al. (2020) investigated the factors influencing trust in robots and artificial intelligence (AI) within trust game scenarios and initial encounters. The study determined if the name and type of opponent, familiarity with robots, robot appearance, and task performance impacted trust levels in robots and AI. To this end, the authors assessed trust in robots and AI using a trust game, where participants decided what portion of $1,000 to share with an opponent. To measure the participant's trust, the shared amount was then tripled, emphasizing the potential mutual gain if the opponent chose to return some of the money; trust was determined by the participant's willingness to risk their own resources for potential reciprocal gains. The researchers utilized online questionnaires to collect demographic data and to measure personality traits (using the 15-item Big Five Inventory), allowing them to determine how these factors influenced trust dynamics during interactions with AI and robots. Data were gathered from a broad cohort of 1,077 participants conducted via Amazon's Mechanical Turk, which offered a demographic mix that spans various ages, genders, and educational backgrounds, ensuring a rich contextual analysis of trust dynamics.

The findings indicated that neither the name of the robot (whether a typical human name or a nickname) nor the type of opponent (robot vs. AI vs. human) significantly impacted the trust participants placed in these entities, contradicting some common assumptions that anthropomorphic features or familiar human names might influence trust levels. Participants with previous exposure to or familiarity with robots reported higher trust levels, suggesting that prior positive experiences or familiarity reduces fear and uncertainty, enhancing trust. The human-like appearance of robots was positively correlated with higher trust levels, such that participants tended to trust robots more when they resembled humans, likely due to the perceived relatability or understandability of their actions. Finally, effective task performance by robots was a strong predictor of trust; participants who observed robots performing tasks effectively and efficiently were more likely to trust them. These findings support the positive impact of familiarity, human-like appearance, and effective task performance on trust in robots and AI. The authors concluded that enhancing the human-like qualities of robots and ensuring their competent performance in interactive tasks can significantly increase trust levels. These results underscore the importance of design and functionality considerations in AI and robot development to align with human expectations and preferences, thereby fostering a trusting relationship.

The notion of personalizing AI to fit diverse user profiles aligns with broader calls for AI systems that are not only technologically proficient but also culturally and psychologically sensitive. To that end, Selezneva et al. (2023) investigated the complex interplay between volitional regulation (i.e., individuals' ability to manage their actions amidst modern world challenges such as technological advances and social instability) and trust dynamics. The study examined how self-trust and trust in others influenced individuals' capacity to initiate and execute personal intentions (i.e., their ability to begin and complete planned actions). This exploration aimed to unravel how trust configurations affect personal agency in managing life's demands and interactions with technology. The authors hypothesized a relationship between various patterns of trust, referred to as trust configurations (i.e., different levels of self-trust and trust in others) and their impact on individuals' effective initiation and completion of intended actions. To examine these dynamics, they used the Self-Confidence Scale (Skripkina, 2014), an adaptation of Rotter's Interpersonal Trust Scale (Dostovalov, 2013), and an adaptation of Kull's Control in Action scale (Shapkin, 1997).

The study involved 210 participants ranging from 27 to 55 years old, with a significant majority being female (72%). Results revealed significant differences in the ability to follow through with intentions, depending on the type of trust relationship: participants with higher self-trust demonstrated greater perceived control over their life choices, while those with lower self-trust encountered difficulties in goal setting and adapting their life circumstances. Their study revealed how different trust configurations (i.e., trust in self and/or others) impact one's ability to act upon personal intentions, which can extend to interactions with AI systems. The study highlights how individuals' trust predispositions influenced their interactions with AI technologies, emphasizing the importance of understanding the various factors that contribute to trust in AI. This understanding forms a framework that examines the different ways people may trust and engage with AI systems, depending on their personal trust configurations. Individuals with robust self-trust may perceive AI as an empowering tool that augments their capabilities and control. In contrast, those with lower self-trust might approach AI with skepticism, impacting their adoption and operational reliance. This distinction enhances our comprehension of the pivotal role trust plays in the human-AI interface, urging the development of AI systems that are finely attuned to the varied psychological landscapes of their users.

Unver and Asan (2022) examined the role of trust in AI-driven healthcare systems and its impact on patient safety, underscoring the increased use of AI in healthcare for various purposes (e.g., diagnosis, drug development, personalized medicine, and patient care monitoring); the authors point out that AI cannot be held liable for flawed decisions in healthcare and does not possess the capacity to be trusted according to traditional definitions of trust. They further discuss the ethical and legal concerns associated with using AI in healthcare, including issues of informed consent, privacy, bias and discrimination, and opacity (black-box problem). Informed consent is essential when using AI health apps and chatbots that collect and analyze data from wearable sensors; privacy is a significant concern, as AI relies on large datasets that may contain sensitive information, and bias and discrimination can occur if AI algorithms are trained on biased or unrepresentative data. AI algorithms' opacity poses additional challenges in understanding AI's decision-making process. The authors suggested that trust in these systems requires all the environmental, psychological, and technological conditions to be responsive to patient safety, further suggesting that trust should be built on a multidimensional relationship between the various actors in the healthcare system, including patients, clinicians, and policymakers. The authors stress the importance of transparency, responsibility, accountability, fairness, and privacy in building trust, concluding that AI has the potential to bring significant benefits to healthcare and patient safety but that trust in these AI systems should be viewed holistically, considering the interdependencies and interactions within the healthcare system.

Bridging the conceptual discourse on trust in AI-driven healthcare systems to practical implications involves examining the nuanced impacts of AI on individual behavior. As ethical considerations evolve into analyses of real-world interactions, it becomes crucial to observe how specific demographics engage with online health information and to examine the dynamic interplay between theoretical constructs of trust and their tangible effects on health information-seeking behaviors. To that end, Hatamlah (2024) conducted a study into the factors influencing adolescents' online health information-seeking behaviors, focusing on trust in online health information, eHealth literacy, parental influence, and the credibility of AI-generated health information. A survey was conducted among 381 adolescents who completed the Trust in Online Health Information (Rowley et al., 2015) eHealth Literacy (Norman & Skinner, 2006), Parental OHIS Behaviors (Hatamlah, 2024), Parental OHIS Mediation (Hatamlah, 2024), and AI-Generated Health Information System Credibility Scores. Additionally, adolescents' eHealth literacy was assessed using a comprehensive eHealth Literacy Assessment. The data were analyzed using Structural Equation Modeling informed by social cognitive theory, and the results confirmed significant relationships between trust in online health information, eHealth literacy, parental behaviors, AI-generated credibility scores, and adolescents' online health information-seeking behaviors.

The results of the study underscore the positive influence of trust in online health information and eHealth literacy on adolescents' disease-related and fitness-related online health information-seeking behaviors. These findings highlight the importance of credible and accessible online resources in supporting effective health management among young individuals. Additionally, the study found that parental behaviors and mediation significantly enhance adolescents' engagement with online health resources, particularly in how they seek information related to disease management and fitness. This reinforcement by parents, together with trustworthy AI-generated health information, underscores the vital role of credible AI applications in promoting informed health decisions among adolescents. Hatamlah's (2024) investigation provides a critical link to understanding trust in AI within healthcare, emphasizing the significance of eHealth literacy and AI-generated content credibility. This study aligns with broader concerns about misinformation and AI hallucinations discussed in AI contexts by highlighting the role of parental influence and AI-generated credibility scores in shaping adolescents' health information-seeking behaviors. It underscores the importance of credible AI applications in enhancing trust, especially relevant to societal reliance on AI technologies and addressing challenges posed by misinformation, thus directly contributing to the discourse on AI as trustworthy social actors.

Transitioning from the exploration of trust within AI-driven healthcare systems, where ethical considerations and patient safety are paramount, to the domain of politics and news, one observes a thematic shift while maintaining a conceptual linkage. In the context of healthcare, the emphasis on trust centers around ethical imperatives, transparency, and accountability. This focus expands as the discussion moves into the sphere of politics and news, where the challenge becomes discerning truth amidst copious information. Here, the dynamics of media consumption, public discourse, and societal trust emerge as pivotal factors in shaping the perceived credibility of AI-generated content by highlighting the necessity for AI systems across domains to exhibit not only technological excellence but also a commitment to ethical responsibility and transparency, directly addressing the challenge of misinformation.

To address that, Sangwon et al. (2020) examined the relationship between media use, public discussion, social trust, and AI news credibility, i.e., the perceived credibility or trustworthiness of news articles or information generated by AI systems or algorithms. The study aimed to provide insights into how media consumption, public discussion, and social trust contribute to the credibility of AI news. The authors hypothesized that media use, specifically through television, social network sites, online news sites, and public discussion, would be positive related to AI news credibility. The authors further hypothesized that social trust would moderate the effect of public discussion on credibility, such that higher levels of trust in others would strengthen the positive relationship between discussion and credibility. The study used a nationwide online survey to collect data. The survey included measures to assess the frequency of media use (i.e., newspaper news use, TV news use, online news site news use, and SNS news use), public discussion, social trust, and AI news credibility. The study drew samples from an online panel registered with Macromill Embrain, one of the largest survey agencies in South Korea. The sample included a total of 1,502 participants who completed the questionnaire. After excluding incomplete responses, the final sample for analysis consisted of 1,294 participants.

The results of the study showed that age, efficacy, TV news use, online news site news use, and SNS news use, together with discussion frequency and social trust, were all significant predictors of AI news credibility. Specifically, age positively influenced AI news credibility, while efficacy demonstrated a strong positive relationship with credibility. Media consumption variables, including TV news use, online news site news use, and SNS news use, all positively correlated with AI news credibility, underscoring the impact of diverse media platforms on perceptions of AI-generated content. Additionally, the frequency of public discussion was positively associated with AI news credibility, highlighting the role of societal engagement in shaping these perceptions. Furthermore, social trust was found to correlate positively with AI news credibility, indicating the foundational importance of trust in social relations for accepting AI-generated news. The analysis also revealed that the interaction between discussion frequency and social trust significantly enhanced AI news credibility, suggesting a moderation effect where social trust strengthens the positive influence of public discourse on credibility perceptions. The insights from Sangwon et al. (2020) significantly contribute to the discourse on trust in AI, particularly in the realm of AI-generated news. Their findings illuminate how media consumption, public dialogue, and social trust combine to mold perceptions of credibility in AI-produced content. Through elucidating the relationship between social trust and the credibility of AI news, these findings echo the need for strategies that not only mitigate misinformation but also enhance the public's discernment abilities in the age of advanced AI technologies.

As the focus shifts from broader societal impacts to detailed mechanics within digital environments, the roles of interpretability and privacy become central to fostering trust in algorithmic news platforms. The interplay between technology design and user perceptions deepens our understanding of how transparency and fairness are pivotal in securing trust in AI systems. Shin et al. (2022) further advanced our understanding by highlighting the critical role of interpretability and privacy in fostering trust in algorithmic news platforms. They examined the relationship between interpretability, trust, privacy, and information disclosure in the context of algorithmic news platforms to determine how users' cognitive processes influence their decisions to disclose personal information in platformized news contexts (i.e., environments where news content is personalized and delivered through AI-driven platforms that aggregate, curate, and tailor information based on user behavior and preferences). The authors proposed several hypotheses, suggesting that interpretability (i.e., the extent to which users can understand the internal mechanics of algorithmic news platforms in human language) influences perceptions of transparency, fairness, and accountability. These enhanced perceptions are predicted to bolster trust in algorithmic news platforms. The authors further hypothesized that trust will positively influence users' privacy concerns and their willingness to disclose personal information, and that privacy will positively influence users' information disclosure in algorithmic news platforms. The study focused on the Naver news recommendation platform in South Korea, utilizing a combination of online ethnographic methods and offline experiments to assess user interactions with algorithm-generated news and video content. Participants recruited through undergraduate and graduate courses related to news media, digital technologies, and consumer research were divided into a control group and an experimental group. The latter was primed with explanatory cues in AI news recommendation systems, aimed at evaluating the impact of interpretability on users' perceptions of transparency, fairness, and accountability, and subsequently, their trust in the platform. This methodological approach was designed to simulate real-world engagement with platformized news, specifically by incorporating user-specific adjustments that algorithmic platforms use to tailor content. These personalization settings, which adjust news delivery based on individual preferences, browsing history, and interactions with content, enhance user engagement by ensuring the news presented is more relevant to each user's interests.

The results demonstrated that interpretability significantly and positively influenced user perceptions of transparency; these enhanced perceptions positively impacted user trust in the platforms, which in turn positively influenced their privacy concerns by reducing apprehensions and positively affected their willingness to disclose personal information. Privacy positively influenced information disclosure by serving as a mediating factor between trust and the willingness to share personal information. This mediation role clarifies how increased trust, bolstered by interpretability, leads to enhanced privacy perceptions, which then encourage greater disclosure of personal information. Additionally, the findings revealed that explanatory cues increased users' understanding of privacy mechanisms, a concept referred to as "privacy heuristics" (i.e., cognitive shortcuts that help users quickly assess privacy risks and benefits, leading to decisions about personal data disclosure). That explanatory cues were positively related to privacy heuristics suggests that clearer explanations increase users' ability to navigate and control their privacy, which in turn positively impacts their willingness to disclose personal information. Results further indicated that interpretability significantly influenced users' perceptions of transparency, fairness, and accountability in algorithmic news platforms, such that interpretability enhanced transparency by making the algorithmic processes clearer to users, which bolstered their trust in the platforms. Furthermore, the clearer explanations provided by interpretability helped users perceive the platforms as fairer and more accountable, positively affecting their trust. These improvements in users' perceptions directly facilitated an increase in trust, underscoring the crucial role of interpretability in building trust through enhanced transparency, fairness, and accountability. The findings suggest that users are more likely to trust algorithmic news platforms when they clearly understand how the algorithms work and how their personal data are used. This has implications for designing and implementing AI systems, as transparency and interpretability can help build trust and mitigate privacy concerns.

After exploring the role of AI in journalism and the broader dynamics of trust, Kuznetsova et al. (2023) examined AI’s function in the arena of political information. This study not only underscores the technological advancements and challenges in AI's ability to manage misinformation but also highlights the societal implications of AI in shaping political discourse. Kuznetsova et al. (2023) evaluated the ability of ChatGPT and Bing Chat (i.e., two large language model (LLM)-based chatbots) to detect the veracity of political information. The authors used a diverse array of topics for evaluation, including COVID-19, Russian aggression against Ukraine, the Holocaust, climate change, and LGBTQ+ debates. The study used AI auditing methodology to determine how the chatbots performed in different languages (English, Russian, and Ukrainian) and how they evaluate statements according to political communication concepts of disinformation, misinformation, and conspiracy theory. The authors developed a codebook with variables such as the accuracy of detecting false, true, and borderline statements, the accuracy of detecting conspiracy theories, and the presence of misinformation or disinformation. Data were collected by manually submitting the prompts to the chatbots and recording the chatbots' responses and were analyzed to identify factors influencing the chatbots' performance in evaluating veracity, detecting conspiracy theories, and identifying misinformation and disinformation. The results showed that ChatGPT performed better than Bing Chat in accurately evaluating statement veracity, with an overall accuracy of 72% across languages without pre-training. Bing Chat had a lower accuracy of 67%. The results highlighted significant disparities in performance between high-resource languages, which have abundant training data, and low-resource languages, which have limited data. This disparity was particularly evident with Ukrainian prompts, where Bing Chat's performance was notably lower due to its classification as a low-resource language with less available training data. The performance of the chatbots showed considerable variations depending on the topics and the sources of the statements. Analysis revealed that chatbots were more accurate in identifying the veracity of statements related to well-documented historical events, such as the Holocaust, where they showed fewer inaccuracies in labeling statements as non-intentionally false. In contrast, statements concerning contemporary political issues, particularly those involving Russian aggression against Ukraine, were more frequently misclassified as intentionally false or were not responded to at all. This variability underscores the influence of the topic's complexity and the source's reliability on the chatbots' ability to assess information accurately.

Progressing from the societal impacts of AI in disseminating news and information to the technical and community-based trust in programming environments underscores the complex and varied nature of trust in AI across different fields. Each domain presents distinct challenges and opportunities, emphasizing the critical importance of cultivating trust, transparency, and ethical considerations in advancing and applying AI technologies. To that end, Cheng et al. (2023) aimed to determine how online communities shape software developers' trust in AI-powered code-generation tools. In particular, the authors sought to understand how online communities help developers build appropriate trust in AI code-generation tools and to investigate the design space for integrating user communities into AI code-generation systems. Researchers also aimed to investigate how online communities shape developers' trust in AI tools and how community features can facilitate appropriate user trust. The study used semi-structured interviews to gather qualitative data from 17 participants who had experience consuming or sharing content about AI code-generation tools in online communities. Results indicated that online communities provided two key aspects that help developers build trust in AI code-generation tools: community-curated experiences and community-generated evaluation signals. Community-curated experiences included vivid descriptions and explanations of others' experiences with AI tools, realistic programming tasks, diverse use cases, and details on project context and dependencies. Community-generated evaluation signals included direct indicators of code quality, context and generation process of code solutions, and identity signals. These evaluation signals helped developers evaluate the quality and trustworthiness of AI-generated code and make informed decisions. The study contributes to understanding how developers build trust in AI-powered code generation tools through their participation in online communities. This study illuminates how online community participation can influence trust development in AI technologies, emphasizing the nuanced interplay between communal feedback and individual trust assessment.

The scrutiny of AI news credibility, influenced by media consumption, public dialogue, and social trust, highlights the pivotal role of trust in AI systems, particularly regarding its broader societal impacts. However, one of the most direct and profound effects is observed within specialized educational settings, where the dynamics of trust and technology converge with pedagogical practices. Amoozadeh et al. (2023) examined the influence of trust in Generative AI (GenAI) on students' engagement and learning outcomes within computer science education. The primary objective was to discern how variations in trust levels among students affect their use of GenAI tools, thereby impacting their motivation, confidence, and proficiency in programming. This study was conducted using a survey administered to 253 students from the University of Houston and the Indian Institute of Technology Kanpur to capture a diverse demographic spectrum, including first-generation and continuing-generation students from undergraduate to graduate levels. The authors hypothesized that students' trust in GenAI would be positively related to their motivation, confidence, and knowledge of programming. To measure these variables, they utilized instruments adapted from existing "trust in AI" tools (Körber, 2019). These instruments included Likert-scale items assessing trust based on the system's output comparability to a competent human and overall reliability. The demographic profile of participants included 210 males, 39 females, and 4 identifying as 'Other', highlighting a significant gender disparity that could influence trust dynamics. The survey explored various aspects of GenAI usage, trust levels, and perceptions of AI tools within programming education. Findings indicated a spectrum of trust levels, where 16% of students expressed distrust, 36% remained neutral, and 47% exhibited trust in GenAI. A positive correlation was found between higher trust levels and increased motivation and confidence in programming for first-generation students. This study contributes to the broader discourse on trust in AI by elucidating its critical role in educational settings, particularly in computer science. The variations in trust, influenced by demographic and educational backgrounds, underscore the need for tailored educational strategies that enhance understanding and mitigate biases inherent in AI systems.

This understanding of trust's influence sets the stage for further exploration into user interactions with AI across different contexts. Hyun et al. (2023) extended this investigation by determining the influence of specific motivations (e.g., information seeking and task efficiency, personalization, social interaction, and playfulness) on user perceptions, particularly feelings of creepiness and trust towards AI technologies like ChatGPT. Four hundred twenty one participants completed instruments adapted from previous research: Information seeking was measured using items from Leung and Wei (1998; Rubin, 1983), task efficiency and social interaction used slightly modified items from Choi and Drumwright (2021), personalization was assessed with items from Baek and Morimoto (2012), playfulness was measured using items from Kim and Baek (2022), creepiness was evaluated with items from Rajaobelina et al. (2021), and trust was gauged using items from Kim and Baek (2022). Results indicated that task efficiency and social interaction unexpectedly increased perceptions of creepiness, contrary to the hypothesized relationships. Conversely, personalization appeared to decrease creepiness and enhance trust, aligning with predictions. These outcomes highlight the dynamics between user motivations and their psychological reactions to AI, suggesting that not all functionalities perceived as efficient or interactive are comforting or trust-enhancing. Their findings suggest that while some user motivations can enhance trust, others might unexpectedly elevate feelings of creepiness, thereby complicating the trust landscape in AI interactions. Understanding these dynamics is essential for designing AI systems that are user-friendly and capable of fostering trust rather than discomfort.

Locus of Control

The studies reviewed so far highlight the need to consider other psychological traits that could moderate interactions between humans and AI. One such trait is locus of control, which refers to an individual's belief system regarding the extent to which they can control events affecting them. This trait is particularly salient in the context of AI, where it could influence user perceptions significantly; individuals with an internal locus of control may perceive AI tools as empowering, enhancing their trust, whereas those with an external locus might view these technologies as threatening or manipulative (Rotter, 1966).

Novozhilova et al. (2024) conducted a study into U.S. public attitudes toward the potential automation of various occupational roles, including therapy, surgical teams, air traffic control, construction sites, news desks, stock investments, and customer service, among others, to examine how individual traits such as locus of control, perceived technological competence, and innovativeness influenced these attitudes. The authors hypothesized that public comfort with AI's role in different occupational domains would be influenced by the likelihood of those domains being automated. Additionally, they predicted that individual traits, such as locus of control, perceived technological competence, and innovativeness, would significantly impact these attitudes. To measure comfort levels with AI managing roles from high-risk sectors like air traffic control to routine tasks like supermarket management, Participants (N = 1150) rated their comfort with AI on a five-point Likert-type scale; to measure individual traits participants completed the 13-item Locus of Control scale (Rotter, 1966) and the 7-item Perceived Technology Competence scale (Katz & Halpern, 2013). The shortened version of the Hurt et al. (1977) scale measured innovativeness and demographic information was also collected. Results revealed that participants with higher incomes, males, and those with greater perceived technological competence demonstrated increased comfort with AI across all domains. Conversely, those with a higher internal locus of control, indicating a preference for personal control over outcomes, showed decreased comfort with AI management. These findings suggest that individual perceptions of control and technological proficiency significantly shape attitudes toward AI in occupational settings. By highlighting the significant role of individual differences in shaping attitudes toward AI, this study provides a vital perspective for future AI implementations, advocating for a user-centric approach that accommodates a wide range of psychological profiles and fosters greater acceptance and trust in AI technologies.

Building upon the insights into individual differences and their impact on AI acceptance, Sharan and Romano (2020) examine the relationship between personality factors (e.g., locus of control) trust in AI vs human advisors. Their study specifically examined how dispositional differences, measured using the Big Five Inventory (John & Srivastava, 1999) and the Internal Control Index (Rotter, 1966) influence individual trust, using decision-making scenarios involving both AI and human advisors. A total of 171 participants engaged in a card-game-based decision-making task designed to measure trust through various behavioral indicators: reaction times to suggestions, agreement with those suggestions (concordance), and self-reported trust ratings. The authors hypothesized that traits such as extraversion, agreeableness, openness, neuroticism, and conscientiousness would influence trust dynamics differently in interactions with AI and human advisors. Moreover, they posited that locus of control would have a significant moderating effect on these relationships, expecting that higher internal locus of control levels would correlate with greater trust in autonomous systems. Results revealed that higher levels of locus of control led to lower trust in both AI and human suggestions, contradicting the initial hypothesis that it would bolster trust. The authors suggested that individuals with a high internal locus of control tend to rely more on their own judgment and experience a greater need to be in control of their decisions, which may explain their reduced trust in external suggestions from both AI and human advisors. Additionally, neuroticism was negatively associated with trust ratings.

Echoing the influence of individual psychological characteristics on technology acceptance, Kraus et al. (2020) examined the mediating potential of emotion (i.e., anxiety) on the relationship between personality traits (i.e., depression, self-esteem, self-efficacy, locus of control) and trust with automation technology. They hypothesized that depressiveness would be negatively related to trust and self-esteem, self-efficacy, and locus of control would be positively related to trust in an automated driving system. Anxiety was expected to mediate these relationships. Participants ($N = 47$) completed the Beck Depression Inventory-II (Beck, 1961), the Locus of Control Scale (Rotter, 1966), the Rosenberg Self-Esteem Scale (Ferring & Filipp, 1996), the General Self-Efficacy Scale Jerusalem & Schwarzer, 1999), State-Trait Anxiety Inventory (Marteau & Bekker, 1992), and the Automation Trust Scale (Jian et al., 2000). The results supported their hypotheses, revealing that higher levels of depressiveness negatively affected trust in automation, while higher self-esteem and self-efficacy were related to increased trust. Specifically, results showed that as depressive symptoms increased, participants reported lower trust levels in automated driving systems, while increases in self-esteem was related to increased trust in these systems. Unexpectedly, locus of control was not related to trust levels in automation. This may indicate that the influence of locus of control on trust might be more complex or context-dependent than previously thought, and it may interact with other factors in ways that were not captured in this specific study setting. Additionally, the mediating role of anxiety was clearly demonstrated, indicating that higher anxiety levels can diminish trust in automated systems. This mediation suggests that anxiety could offset the positive impacts of high self-esteem and self-efficacy on trust in automation, particularly under conditions of uncertainty or perceived risk associated with automated technologies. The study by Kraus et al. (2020) advances our understanding of how psychological traits and anxiety influence trust in automated technologies, emphasizing the importance of considering both stable personality traits and transient emotional states in AI design.

Collectively, these studies form a coherent narrative that underscores the multifaceted nature of trust in AI, highlighting the pivotal role of locus of control alongside other psychological traits. This nuanced understanding remains critical for developing AI systems that are not only technologically advanced but also psychologically attuned to the diverse needs of users. Insights from these studies can help inform future research and potentially guide AI developers in enhancing the efficacy and acceptance of AI technologies, ensuring they align more closely with user expectations and psychological profiles.

Self-Efficacy

Exploring the psychological underpinnings that govern technology use, the subsequent study by Compeau and Higgins shifts focus to internal belief systems, specifically computer self-efficacy, that enable users to harness technology effectively. This progression underscores a shift from external attributes of robots and AI that inspire trust to the internal attributes of users that facilitate competent technology engagement.

Compeau and Higgins (1995) examined the role of computer self-efficacy in enhancing computer usage. The study aimed to develop and further validate the newly developed 10-item measure of computer self-efficacy and examine its impact on users' expectations, emotional reactions, and actual computer use. The authors hypothesized that social encouragement, organizational support, and observing others use computers would positively influence individuals' self-efficacy and subsequent interactions with technology. Participants ($N = 2{,}100$) were managers and professionals, predominantly male (83%) who completed the 10-item Computer Self-Efficacy Scale (developed by the authors) to assess beliefs about their capabilities to use computers effectively. Additionally, they used the Encouragement by Others Scale (e.g., "My colleagues encourage me to use computers"; Burhardt & Brass, 1990) and the Organizational Support Scale (e.g., "My organization provides adequate resources for computer training"; Thompson et al., 1991). The survey was conducted through a combination of pilot (100 participants) and main studies, capturing a comprehensive view of the role of computer self-efficacy across different organizational contexts.

As predicted, results indicated positive relationships between encouragement from others and organizational support with computer self-efficacy, positively affecting user emotions and negatively correlated with user anxiety. Using hierarchical linear regression models, the findings revealed significant positive correlations between supportive factors (i.e., encouragement from others and organizational support) and individuals' computer self-efficacy. Increased self-efficacy was significantly and positively associated with enhanced positive emotions towards computer use and a reduction in anxiety. These changes in emotions and anxiety were linked to more frequent and effective computer use. The supportive factors accounted for a significant portion of the variance in computer self-efficacy, highlighting their critical role in reducing anxiety and boosting user affect, thereby fostering better computer usage habits. These findings highlighted the critical role of environmental and social support in enhancing self-efficacy, directly impacting users' emotional experiences and behavior towards technology.

While Compeau and Higgins (1995) focused on developing and validating a measure of computer self-efficacy and its impact on usage, Torkzadeh et al. (2003) aimed to deepen the understanding of this construct by testing a 4-factor model of computer self-efficacy. To that end, Torkzadeh et al. examined the structure of computer self-efficacy by examining the relationship between individuals' self-perceptions of their computer-related capabilities and their engagement and performance with technology. This model included dimensions such as beginning skills, file and software management skills, advanced skills, and mainframe skills, specifically tailored to reflect the progressive complexity of computer-related tasks. They employed the revised Computer Self-Efficacy Scale (CSES) originally developed by Murphy et al. (1989), a self-report questionnaire that assesses individuals' perceptions of their capability in specific computer-related knowledge and skills areas, including beginning skills, file and software skills, advanced skills, and mainframe skills. The study was conducted with 414 undergraduate business students, evenly distributed between 202 male and 212 female students, enrolled in an introductory computer course. The findings revealed that beginning skills were positively correlated with file and software skills, advanced skills, and mainframe skills; file and software skills were positively related to advanced skills and mainframe skills; advanced skills were positively related to mainframe skills. These positive associations indicated a cohesive structure of the model and validated its multidimensionality, showing that as self-efficacy in one area increases, so does self-efficacy in the others. These findings underscore the importance of designing educational programs that specifically aim to bolster self-efficacy, which is critical for fostering effective technology use and enhancing trust in AI.

Montag et al. (2023) investigated the relationships among technology self-efficacy, trust in (automated) technology, and their impacts on the fear and acceptance of AI, hypothesizing that trust in technology would mediate the relationship between self-efficacy and acceptance (or fear) of AI. To test these hypotheses, the researchers employed several well-established instruments: the Technology Self-Efficacy Scale (Holcomb et al., 2004), the Propensity to Trust in (Automated) Technology Scale (Merritt et al., 2013), and the Attitudes Towards Artificial Intelligence Scale (Sindermann, 2021). The sample comprised 289 participants, ranging in age from 18 to 70, with a mean age of 29.26 years. The demographic makeup was notably skewed, consisting of 73 males and 216 females, a disparity that the researchers acknowledged could influence the generalizability of the findings. Education levels varied, providing a broad insight into societal attitudes towards AI. Their findings revealed significant positive correlations between technology self-efficacy, trust in (automated) technology, and acceptance of AI, suggesting that individuals who feel more capable in their interactions with technology are also more likely to trust and accept AI systems. Conversely, there were significant negative correlations between technology self-efficacy and fear of AI, suggesting that as individuals' confidence in their technological abilities increases, their fear of AI decreases. The mediation analysis further clarified these relationships, showing that trust in technology plays a crucial mediating role. Specifically, the mediation models revealed that the effect of technology self-efficacy on acceptance of AI was partially mediated by trust in technology, meaning that higher self-efficacy in technology not only directly contributes to greater acceptance of AI but also does so indirectly by fostering greater trust in technological systems. The same mediating effect of trust was observed in relation to fear of AI, with higher technology self-efficacy reducing fear through enhanced trust.

To extend the role self-efficacy plays in technology interaction, Choi (2021) investigated the acceptance of AI technology, focusing on the factors influencing employee acceptance of AI technology within service sectors. The authors hypothesized that role clarity, motivation, user ability (i.e., self-efficacy), privacy concerns, and trust would significantly influence employees' willingness to integrate AI into their workflows. Specifically, it was proposed that clear role definitions, high motivation, and strong user ability would positively affect AI acceptance, while privacy concerns would negatively impact acceptance and trust would enhance the positive effects of user ability on acceptance. The study utilized a detailed survey methodology, incorporating scales specifically adapted for measuring role clarity, motivation, user ability, privacy concerns, and trust. These included Korean employees (*N* = 454) recruited through a online survey platform completed the Role Conflict and Ambiguity Scale (Rizzo et al., 1970) and the Trust in Technology Scale (Jarvenpaa et al., 1999). As predicted, findings revealed positive correlations, with role clarity, motivation, and user ability significantly increasing employees' willingness to accept AI. Privacy concerns negatively moderated the relationship between role clarity and acceptance, by decreasing acceptance levels when privacy concerns were high. Conversely, trust significantly moderated, by strengthening, the positive influence of user ability on acceptance. These results underscore the importance of clear role definitions, motivational factors, and trust-building measures in AI deployment strategies within organizations. Such findings are critical for AI developers and organizational leaders aiming to implement AI technologies that align with employee expectations and adhere to ethical standards, enhancing overall trust in AI systems.

Collectively, these studies form a coherent narrative that underscores the multifaceted nature of trust in AI, highlighting the pivotal role of locus of control alongside other psychological traits such as self-efficacy, privacy concerns, and trust dynamics. This nuanced understanding remains essential for developing AI systems that are not only technologically advanced but also psychologically attuned to the diverse needs of users. By incorporating these insights into the design and implementation of AI, researchers and developers can enhance the efficacy and acceptance of AI technologies, ensuring they align more closely with user expectations and psychological profiles. Chapter 3 will present the research design, sampling strategy, methodology, recruitment, and data collection procedures. It will also discuss the instrumentation, data analysis plan, threats to validity, and ethical considerations, as they apply to the proposed study.

Chapter 3: Research Method

The purpose of the proposed study was to determine the extent to which locus of control moderates the relationship between self-efficacy and trust in AI among English-speaking adults between 22 and 55 living in the United States. Chapter 3 includes a detailed discussion of the sample, sampling methods, recruitment procedures, inclusion and exclusion criteria, and data collection procedures. Instruments to measure variables are discussed in terms of their validity, reliability, and justification for use in this study. Additionally, I will outline the data analysis plan, including multiple regression. The chapter concludes with a discussion of potential threats to validity and ethical considerations related to participant recruitment, data collection, and analysis.

Research Design and Rationale

I employed a cross-sectional quantitative survey design to determine the extent to which locus of control moderates the relationship between self-efficacy and trust in AI among English-speaking adults between 22 and 55 years living in the United States. The cross-sectional design was appropriate, as data were gathered at a single point in time. A quantitative survey design was the most suitable approach for this research, as it facilitated analysis of relationships between variables of interest. Surveys were distributed online via MTurk, which is an efficient means for reaching a broad and diverse participant base. This method was chosen due to time and cost advantages which are typical of online survey distribution methods. Using MTurk allowed for inclusion of a diverse population, enhancing external validity of findings by ensuring representation beyond specialized groups.

Methodology

Population

The target population for this study was English-speaking adults between 22 and 55 living in the United States. This age range was selected to ensure participants were within the working-age population and were therefore likely to have regular interactions with AI technologies in various contexts, such as in the workplace or through consumer applications. For the purposes of this study, the population included individuals from diverse educational and professional backgrounds, ensuring a broad representation of experiences with AI.

According to the U.S. Census Bureau (2023), the population of adults between 22 and 55 encompasses a significant portion of the working age demographic, with over 100 million individuals. Of these, approximately 78.5% are English speakers or speak English fluently, narrowing the target population to approximately 78.5 million adults (U.S. Census Bureau, 2021).

With the increasing integration of AI into everyday life, a significant portion of U.S. adults interact with AI-driven technologies. For instance, 47% of Americans report regularly using AI-powered digital voice assistants such as Siri and Alexa primarily on smartphones and smart home devices (Olmstead, 2017). AI is also making strides in terms of customer service where 91.9% of users have interacted with AI-based customer service systems such as chatbots; 88.5% noted AI in this area has become common (Chaturvedi et al., 2023). In the workplace, over 50% of U.S. companies have adopted AI tools to improve productivity, with sectors such as technology and finance most prominent (Eastwood, 2022). These advancements reflect a growing reliance on AI technologies across both personal and professional contexts. As a result, approximately 62.97% of English-speaking American adults regularly interact with AI technologies in both personal and professional contexts. This brings the estimated total population for this study to approximately 49.4 million. This large and diverse population base was necessary to capture a wide range of perspectives on self-efficacy, locus of control, and trust in AI, providing a robust foundation for analysis.

Sampling and Sampling Procedures

A convenience sampling strategy was employed for this study. Participants were recruited using MTurk. To be included in this study, participants were required to be between 22 and 55 and English-speaking adults living in the United States. I excluded individuals under the age of 22 or over 55, non-English speakers, and individuals residing outside the United States. This was to ensure consistency in terms of language proficiency and cultural background regarding AI use and trust.

To determine the minimum sample size for the study, a power analysis using G*Power 3.1 was completed. The minimum recommended sample size was calculated using an alpha level of 0.05, power of 0.80, and estimated effect size of 0.15 for the moderator. Analysis included one tested predictor (interaction variable) and three total predictor variables (locus of control, self-efficacy, and the interaction variable). The recommended sample size as a result of these parameters was 89 participants. To account for potential participant dropout or incomplete responses, oversampling was conducted to ensure sufficient data for analysis. Additionally, data cleaning procedures were used to remove cases with incomplete data.

The estimated effect size was chosen based on previous empirical research examining relationships between psychological traits and trust in technology. Studies exploring the role of self-efficacy and locus of control in terms of shaping trust in technology reported small to medium effect sizes. Rotter (1966) demonstrated locus of control significantly influences trust in external systems, with effect sizes typically ranging from small to medium (rs = .20 to .30). Similarly, Bandura (1977) stated self-efficacy had a positive impact on confidence in technology use. Moderation effects in psychological studies, particularly those involving individual traits such as self-efficacy and locus of control, often yield small to moderate effect sizes (Aguinis et al., 2005). Therefore, an estimated effect size of 0.15 for the moderation effect in this study was consistent with existing literature in the field and provided a robust basis for sample size determination.

Procedures for Recruitment, Participation, and Data Collection

With IRB approval, participants were recruited using MTurk. MTurk provided a diverse participant pool which was used to recruit English-speaking adults between 22 and 55 living in the United States. Recruitment was facilitated through MTurk's built-in filtering system to ensure only participants who met eligibility criteria could participate.

To participate in the study, individuals were first presented with informed consent forms directly via MTurk. This consent form outlined the purpose of the study, privacy-related concerns, the voluntary nature of participation, and risks and benefits. To maintain participant anonymity, participants were required to indicate their agreement. Responses were collected anonymously, ensuring no identifying information was required or recorded. Those who selected the disagree option or did not meet inclusion criteria were automatically redirected to a thank you page, where they were thanked for their time and interest. Participants who consented and met eligibility criteria then proceeded to the survey, which included demographic questions regarding age, gender, race, education level, and employment status. The survey also included self-report instruments to measure locus of control, self-efficacy, and trust in AI. The survey was expected to take approximately 25 minutes to complete.

At the conclusion of the survey, participants were directed to a debriefing page that provided additional details about the study, clarified any potential misunderstandings, and offered contact information for mental health resources in case participants experienced discomfort during or after the study. Participants were also provided with Walden University's participant advocate should they have any questions or if they wished to receive a summary of the study's results once the study was completed.

Instrumentation and Operationalization of Constructs

Demographic Questionnaire

A demographic questionnaire was administered to gather basic participant information and verify eligibility, including participants' age, gender, race, education level, employment status, and state of residence. In addition to these basic demographic questions, the survey also asked participants to provide information on their experience with AI technologies, such as the types of AI systems they use (e.g., personal assistants, AI-powered customer service tools, workplace tools), the frequency of their interactions, and the primary contexts in which they use AI (personal or professional). The demographic questionnaire was expected to take approximately five minutes to complete.

Digital Technology Self-Efficacy Scale (DTSES)

The Digital Technology Self-Efficacy Scale (DTSES), adapted by Hughes (2013) from the Computer Self-Efficacy Scale originally developed by Cassidy and Eachus (2002), was used to measure participants' confidence in their ability to engage with and effectively use digital technology. The DTSES consists of 17 items, with responses rated on a 4-point Likert scale ranging from 1 (*strongly disagree*) to 4 (*strongly agree*). Several items are reverse-coded to ensure consistency in scoring direction. Sample items include statements such as "I find working with digital technology very easy" and "I often have difficulties when trying to learn how to use a new software package or online application." The DTSES is scored by summing individual responses, yielding total scores from 17 to 68. Higher scores indicate greater digital technology self-efficacy. The DTSES has been effectively used in research exploring confidence and attitudes toward digital technology, making it relevant for assessing user interactions with AI-driven systems. Completion of the DTSES typically takes approximately seven minutes. This scale is publicly available for research use without author permission (Hughes, 2013).

Previous research using the DTSES has demonstrated strong psychometric properties, including high internal consistency reliability and robust construct validity. Hughes (2013) reported high reliability, with Cronbach's alpha coefficients ranging from

.956 to .960 across three distinct samples, indicating excellent internal consistency. Similarly, Ok, Hughes, and Lee (2017) found comparable internal consistency, reporting Cronbach's alpha values between .941 and .965 in separate samples, further supporting the reliability of the DTSES. Construct validity was established through exploratory and confirmatory factor analyses, indicating a clear and stable factor structure across multiple studies (Holcomb et al., 2004; Hughes, 2013). Additionally, Hughes (2013) demonstrated good convergent validity through positive correlations between DTSES scores and related measures of technology integration confidence, further supporting its validity as a measure of digital technology self-efficacy.

Rotter's Internal-External Locus of Control Scale

Rotter's Internal-External Locus of Control Scale (Rotter, 1966) was used to measure participants' beliefs about the control they have over the outcomes of their actions. Given the increasing use of AI in decision-making, understanding whether individuals attribute outcomes to their own actions (internal locus) or external factors (external locus) is particularly important. Individuals with a more internal locus of control may feel empowered by AI tools, while those with an external locus may view them as unpredictable or even threatening (Rotter, 1966). The scale consists of 23 forced-choice items, where participants must choose between two opposing statements reflecting either an internal or external locus of control. Sample items include statements such as, "I believe my success depends mostly on my effort" (internal) vs. "Luck plays a big part in my success" (external). Participants' responses were summed to produce an overall score, with higher scores indicating a more external locus of control and lower scores reflecting a more internal locus of control. This scale took approximately 8 minutes to complete. Rotter’s Internal-External Locus of Control Scale is in the public domain and may be used for research purposes without author permission.

Rotter’s Internal-External Locus of Control Scale has been widely used in psychological research for decades, showing strong reliability and validity. The scale has demonstrated high internal consistency, with Cronbach’s alpha values typically ranging from .74 to .85 across different populations and research settings. (Beretvas et al., 2008; Lefcourt, 1991). Studies have also shown high test-retest reliability over extended periods, with reported correlations of r = .70 to r = .75 over a one-year period (Rotter, 1966).

In terms of validity, Rotter's scale has consistently shown convergent validity with related constructs such as self-efficacy and self-esteem, correlating with measures of self-efficacy in the range of r = .30 to .50 (Lefcourt, 1991). The scale has also been used successfully in a variety of cultural contexts, further supporting its robustness as a measure of locus of control (Berry et al., 1992; Furnham & Steele, 1993). Although Rotter's scale was developed in the 1960s, it has been successfully applied in modern contexts, including studies on technology adoption and trust in automation (Sharan & Romano, 2020; Taffesse & Tadesse, 2017; Hsia et al., 2014).

Trust in Automation Scale

The Trust in Automation Scale (Jian et al., 2000) was used to measure participants' general trust in automation systems. The original scale was developed to assess cognitive trust (beliefs about the competence and reliability of automation) and affective trust (emotional comfort and willingness to rely on automated systems). This distinction is important, as users may cognitively trust AI but still feel hesitant to rely on it in critical situations (Jian et al., 2000). The scale comprises 12 items designed to capture participants' attitudes toward automation reliability, functionality, and predictability. Example items include statements such as "I trust the system to make decisions without supervision" and "I feel comfortable relying on the system for important tasks." Participants rated their agreement with each statement on a 7-point Likert scale ranging from 1 (not at all) to 7 (extremely). Scores were summed across items, with higher scores indicating greater trust in automated systems. Completion of this scale took approximately five minutes. The Trust in Automation Scale is publicly available for research use without author permission (Jian et al., 2000).

Given that this study specifically focused on AI systems, participants were explicitly instructed prior to completing the Trust in Automation Scale that "the system" referred to AI tools and systems. Although examples of AI (e.g., voice assistants, chatbots, recommendation systems, productivity tools, autonomous systems) were provided earlier in the survey during demographic questions, they were not explicitly repeated within the trust items themselves. This approach posed a limitation, as participants' interpretations of "AI" may have varied depending on their personal experience and understanding, potentially influencing their responses and affecting the precision of trust measurement

The Trust in Automation Scale has demonstrated high internal consistency, with Cronbach's alpha values ranging from .88 to .92 in previous research (Jian et al., 2000).

Additionally, the scale has shown strong test-retest reliability, with a correlation of r = .84 over a 2-week period. Reliability has been demonstrated across various domains, including healthcare, aviation, and autonomous vehicle research (Jian et al., 2000; Lee & See, 2004). For example, Lee and See (2004) highlighted its use in assessing trust in automated systems in high-stakes environments, such as aviation and autonomous vehicles.

The scale demonstrates good construct validity, with Jian et al. (2000) reporting significant correlations with related measures, including the Human-Robot Interaction Trust Scale (r = .68) and trust constructs from the Technology Acceptance Model (r = .64). Discriminant validity was supported by weak correlations with theoretically unrelated constructs, such as anxiety and general attitudes toward technology (r < .30; Jian et al., 2000).

Data Analysis Plan

All data were downloaded from MTurk platform directly into SPSS version 29.0 for analysis. Multiple regression with moderation was performed to assess the relationships among locus of control, self-efficacy, and trust in AI, as well as to test whether locus of control moderates the relationship between self-efficacy and trust in AI. An interaction variable was created by multiplying the two mean-centered independent variables (i.e., locus of control × self-efficacy) to test the moderation effect.

The dependent variable (trust in AI) was regressed on the independent variables (locus of control and self-efficacy) and the interaction variable (locus of control × self-efficacy). The regression model was built by entering all predictors simultaneously, including the interaction term, to explicitly test the moderation effect. All assumptions for multiple regression were carefully checked using SPSS to ensure the validity of the analysis. The assumption of normality was examined by inspecting histograms and Q-Q plots of the regression standardized residuals, supplemented by assessing skewness and kurtosis values. Additionally, homoscedasticity was evaluated by examining scatterplots of residuals versus predicted values, with evenly dispersed residuals indicating the assumption was met. The assumption of independence of residuals was assessed using the Durbin-Watson statistic; a Durbin-Watson value close to 2.0 indicated no autocorrelation. Linearity was confirmed through visual inspection of the normal P-P plot of standardized residuals and scatterplots, confirming a linear relationship between observed and predicted values.

To address the assumption of multicollinearity, Variance Inflation Factor (VIF) values were calculated for all predictors. VIF values exceeding 10 would have indicated a potential multicollinearity problem. While multicollinearity is often a concern when including interaction terms, Friedrich (1982) suggested that multicollinearity arising from interaction terms may not present significant issues in certain moderation analyses, and the model's results were interpreted accordingly. In addition to checking these assumptions, effect sizes and confidence intervals were calculated to provide insight into the strength and precision of the relationships examined. All statistical analyses were conducted with a significance level of $p < .05$, and any potential violations of assumptions were carefully reviewed to ensure the robustness of the findings.

Research Questions

RQ1: Does locus of control predict trust in AI?

$H_0$1: Locus of control does not predict trust in AI.

H_a1: Locus of control does predict trust in AI.

RQ2: Does self-efficacy predict trust in AI?

$H_0$2: Self-efficacy does not predict trust in AI.

H_a2: Self-efficacy does predict trust in AI.

RQ3: Does locus of control moderate the relationship between self-efficacy and trust in AI?

$H_0$3: Locus of control does not moderate the relationship between self-efficacy and trust in AI.

H_a3: Locus of control does moderate the relationship between self-efficacy and trust in AI.

Threats to Validity

Several potential threats to validity were associated with the design and sampling method used in this study. The first potential threat was related to the sampling strategy. With a convenience sampling method via Amazon Mechanical Turk (MTurk), there was a risk of selection bias, as participants who chose to respond might differ significantly from those who did not. This could limit the generalizability of the findings to the broader population. To reduce the impact of selection bias, demographic screening criteria were applied to ensure that the sample included a diverse range of participants in terms of age, gender, educational background, and geographic location. Additionally, MTurk has been recognized for providing a more diverse and representative sample than traditional convenience sampling methods (Buhrmester et al., 2016), which further mitigated the risk of bias.

Another threat to validity was nonresponse bias, which could occur if participants did not complete the entire survey or dropped out partway through. Nonresponse bias was a concern because it could indicate systematic differences between those who completed the survey and those who did not. For instance, individuals with lower levels of trust in AI or who were less confident in using technology might have been more likely to abandon the survey. To minimize this threat, the survey was structured in a way that encouraged completion, with participants reminded that the study was anonymous. Additionally, participants were unable to skip questions, reducing the risk of missing data due to incomplete responses. If a participant chose not to complete the survey, their data were excluded from the analysis.

There was also a potential risk of social desirability bias, where participants might provide responses that they believed were more socially acceptable or favorable rather than being fully honest. This was especially relevant when assessing self-efficacy and trust in AI, as participants might have felt pressure to portray themselves as more confident or trusting of AI technologies than they truly were. To mitigate this risk, participants were reminded multiple times throughout the survey that their responses were anonymous, and that no identifying information would be collected. This anonymity was expected to encourage more honest and candid responses, thereby reducing the likelihood of social desirability bias.

Lastly, the use of self-reported measures introduced a risk of response bias, where participants' subjective interpretations of the questions might influence the accuracy of the data. Although this could not be eliminated entirely, using well-validated instruments such as the Digital Technology Self-Efficacy Scale, Rotter’s Internal-External Locus of Control Scale, and the Trust in Automation Scale helped ensure that the measures were reliable and valid. Additionally, instructions were clearly stated to reduce potential confusion among participants.

Ethical Procedures

Before any data collection began, approval was obtained from Walden University’s Institutional Review Board (IRB) to ensure that the study met all ethical guidelines. Participants were provided with an informed consent form, offering a brief description of the study, including its purpose, potential risks and benefits, and their privacy rights. Participation in the study was voluntary, and participants were informed that they could withdraw at any time without penalty. The survey was administered anonymously via Amazon Mechanical Turk (MTurk); no identifying information was collected at any point, thereby minimizing any risks to participant privacy. Given that this study involved the assessment of locus of control, self-efficacy, and trust in AI, no foreseeable psychological or emotional risks were anticipated. However, to ensure participants' comfort, they were reminded in the consent form that they could discontinue the survey at any point if they felt uncomfortable.

In terms of data security, all survey responses were encrypted and stored on MTurk’s secure platform until they were transferred to SPSS for data analysis. Once the data were downloaded, they were stored on a password-protected computer accessible only by the researcher. The data will be retained for five years following the completion of the study, after which they will be securely deleted in accordance with Walden University's data retention policy. Participants were also assured of the anonymity of their responses. No identifying information was linked to the data at any point, ensuring complete anonymity. Additionally, all participants were given clear instructions in the informed consent form on how to contact me or Walden University's participant advocate if they had any questions or concerns about the study. Overall, these procedures were designed to safeguard participant privacy and ensure ethical treatment throughout the research process.

Summary

The purpose of this study was to address a gap in literature by determining the extent to which locus of control moderated the relationship between self-efficacy and trust in AI among English-speaking adults between 22 and 55 living in the United States. I employed a quantitative correlational cross-sectional survey design using MTurk for participant recruitment and administration of surveys. To be included in the study, participants had to be English-speaking adults within the specified age range and residing in the United States. Participants used Rotter's Internal-External Locus of Control Scale, the DTSES, and the Trust in Automation Scale. Data were analyzed using multiple regression with moderation to examine the moderating influence of locus of control on the relationship between self-efficacy and trust in AI. All assumptions of multiple regression were thoroughly checked for compliance.

Potential threats to validity, such as selection bias, nonresponse bias, and social desirability bias, were identified, and strategies were outlined to mitigate these risks. Ethical considerations were thoroughly addressed, including participant privacy, anonymity, and data security. IRB approval was obtained before any data collection took place. Data will be stored securely and deleted after 5 years in compliance with data retention policies.

Chapter 4 includes results of the study, including descriptive statistics, regression analyses, and a discussion of findings in relation to research questions and hypotheses.

Chapter 4: Results

The purpose of this quantitative study was to determine the extent to which locus of control moderates the relationship between self-efficacy and trust in AI among English-speaking adults between 22 and 55 living in the United States. I used multiple regression with moderation to address this topic.

This chapter includes results of statistical analyses that were conducted to address research questions, beginning with an overview of the data collection process and description of the sample, including recruitment, response rates, and demographic characteristics. Next, I address descriptive statistics involving study variables, followed by assumption testing for multiple regression, including assessments of normality, linearity, homoscedasticity, and multicollinearity. Results of multiple regression analyses to test hypotheses are reported. The chapter concludes with a summary of findings.

Research Questions and Hypotheses

RQ1: Does locus of control predict trust in AI?

H_{o1}: Locus of control does not predict trust in AI.

H_{a1}: Locus of control does predict trust in AI.

RQ2: Does self-efficacy predict trust in AI?

H_{o2}: Self-efficacy does not predict trust in AI.

$H_{a2:}$ Self-efficacy does predict trust in AI.

RQ3: Does locus of control moderate the relationship between self-efficacy and trust in AI?

H_{o3}: Locus of control does not moderate the relationship between self-efficacy and trust in AI.

H_{a3}: Locus of control does moderate the relationship between self-efficacy and trust in AI.

Data Collection

Initial recruitment was conducted using MTurk. All participants were required to be English-speaking adults between 22 and 55 residing in the United States. MTurk's built-in screening tools were used to ensure only eligible participants could access surveys. Individuals who did not meet criteria were automatically excluded. While G*Power 3.1 determined that a minimum sample size of 89 participants was required, oversampling was conducted to account for potential incomplete or invalid responses and participant dropout. Data collection was completed within 1 week. The anonymous online survey began with an informed consent form which described the purpose of the study, participant requirements, confidentiality measures, voluntary nature of participation, potential risks and benefits, and contact information for both me and Walden University's participant advocate. Participants who consented to participate proceeded to complete surveys, while those who declined exited the platform.

Following the informed consent form, participants were asked demographic questions regarding their age, gender, race/ethnicity, education level, employment status, household income, disability status, frequency and context of AI interaction, and primary types of AI systems they used. The survey also included Rotter's Internal-External Locus of Control Scale, the DTSES, and Trust in Automation Scale. The estimated time to complete the survey was 25 minutes. A total of 125 participants initially completed the survey. Multivariate outlier analysis using Mahalanobis distance led to identifying seven cases that were subsequently removed to improve data integrity and meet assumptions of multiple regression, resulting in a final sample size of 118 participants.

Of these participants, 82.2% identified as male and 17.8% as female. All participants were between 22 and 55, which was consistent with eligibility requirements. The majority identified as White (86.4%), followed by Asian (9.3%), Black or African American (2.5%), and Native American or Alaska Native (1.7%). Most participants had completed bachelor's degrees (64.4%), followed by master's degrees (32.2%). Employment status was predominantly full-time (94.1%), with smaller proportions of part-time, self-employed, unemployed, and students. The most frequently reported household income range was $75,000 to $99,999 (46.6%). Approximately 48.3% reported no disability, while 27.1% indicated a physical disability, 11.9% a sensory disability, 7.6% a cognitive or learning disability, and 8.5% a mental health condition (see Table 1).

Table 1

Participant Demographics

Variable	***n***	**%**
Age		
22-25	3	2.5%
26-30	18	15.3%
31-35	19	16.1%
36-40	25	21.2%
41-45	34	28.8%

46-50	16	13.6%
51-55	3	2.5%
Gender		
Male	97	82.2%
Female	21	17.8%
Race/Ethnicity		
White	102	86.4%
Black or African American	3	2.5%
Asian	11	9.3%
Native American or Alaska Native	2	1.7%
Education Level		
Some college	3	2.5%
Associate's degree	1	0.8%
Bachelor's degree	76	64.4%
Master's degree	38	32.2%
Employment Status		
Employed full-time	111	94.1%
Employed part-time	1	0.8%
Self-employed	3	2.5%
Unemployed	1	0.8%
Student	2	1.7%
Annual Household Income		
Less than $25,000	10	8.5%
$25,000 - $49,999	18	15.3%
$50,000 - $74,999	26	22.0%
$75,000 - $99,999	55	46.6%
$100,000 - $149,999	5	4.2%
$150,000 or more	4	3.4%
Disability Status		
Physical disability	32	27.1%
Sensory disability	14	11.9%
Cognitive/learning disability	9	7.6%
Mental health condition	10	8.5%
No disability	57	48.3%

As for participants' AI interaction characteristics, 53.4% reported interacting with AI systems daily, while 31.4% interacted weekly and 12.7% interacted monthly. The most common AI systems were chatbots such as ChatGPT or customer service bots, which were mentioned by 55.9% of participants. Voice assistants such as Alexa or Siri were the second most reported at 20.3%, followed by recommendation systems like Netflix or Spotify (14.4%). The largest proportion of participants indicated using AI in professional contexts (42.4%), while 28.8% reported using AI in personal contexts and another 28.8% reported using AI in both personal and professional contexts. Regarding concerns regarding AI technologies, the most frequently reported concerns were privacy (33.1%) and accuracy (33.1%), as well as concerns related to bias, accountability, safety, and lack of transparency (see Table 2).

Table 2

Participant AI Interaction Characteristics

Variable	*n*	%
AI Interaction Frequency		
Daily	63	53.4%
Weekly	37	31.4%
Monthly	15	12.7%
Rarely	3	2.5%
Primary AI System Type		
Voice assistants	24	20.3%
Chatbots	66	55.9%
Recommendation systems	17	14.4%
Productivity tools	9	7.6%
Autonomous systems	2	1.7%
Primary AI Use Context		
Personal	34	28.8%
Professional	50	42.4%
Both	34	28.8%

Because I used convenience sampling strategy rather than random sampling, sample characteristics may not be fully representative of the broader U.S. population. Therefore, generalization of results is limited to individuals who participated in MTurk surveys and met specified inclusion criteria. The final dataset was exported for statistical analysis using SPSS Version 29.

Results

Descriptive Statistics

Means, standard deviations, and bivariate correlations for all study variables are presented in Table 3. Digital technology self-efficacy had a mean score of 44.15 (*SD* = 5.45), trust in automation had a mean score of 62.16 (*SD* = 12.93), and locus of control had a mean score of 34.00 (*SD* = 4.89). Bivariate correlations indicated that digital technology self-efficacy and trust in automation were significantly and positively correlated ($r(118) = .238$, $p = .005$), suggesting that higher self-efficacy was associated with greater trust in automation. No significant relationship was observed between locus of control and trust in automation ($r(118) = -.086$, $p = .177$), nor between digital technology self-efficacy and locus of control ($r(118) = -.142$, $p = .063$).

Skewness and kurtosis values for all variables fell within acceptable ranges (±1), indicating approximate normality. Although the Shapiro-Wilk tests for digital technology self-efficacy ($W(118) = 0.970$, $p = .01$) and trust in automation ($W(118) = 0.925$, $p < .001$) were statistically significant, visual inspection of histograms and Q-Q plots supported the assumption of normality. Locus of control demonstrated a Shapiro-Wilk value of $W(118) = 0.980$, $p = .074$, further supporting approximate normality.

Table 3

Means, Standard Deviations, and Normality Statistics for Study Variables (N = 125)

Variable	*Min*	*Max*	*Mean*	*SD*	*Skewness*	*Kurtosis*

Digital Technology Self-Efficacy	27	55	44.15	5.45	-0.616	0.439
Trust in Automation	30	81	62.16	12.93	-0.779	-0.112
Locus of Control	23	46	34.00	4.89	-0.322	-0.144

Exploratory Data Analysis

Normality

Univariate outliers were evaluated by inspecting standardized residuals for each variable. For all variables, standardized residuals ranged from -2.483 to 1.978, indicating no univariate outliers were present in the dataset (Warner, 2013). Tests of normality were performed using the Kolmogorov-Smirnov and Shapiro-Wilk tests. Results revealed statistically significant findings for some variables. For digital technology self-efficacy, the Kolmogorov-Smirnov statistic was $D(118) = 0.072$, $p = .194$, and the Shapiro-Wilk statistic was $W(118) = 0.970$, $p = .010$. Trust in automation yielded a Kolmogorov-Smirnov value of $D(118) = 0.139$, $p < .001$, and a Shapiro-Wilk value of $W(118) = 0.925$, $p < .001$. Locus of control had a Kolmogorov-Smirnov value of $D(118) = 0.085$, $p = .037$, and a Shapiro-Wilk value of $W(118) = 0.980$, $p = .074$. Although these results indicate deviations from normality, the moderate-to-large sample size ($N = 118$) mitigates the impact of minor deviations and allows for robust statistical analyses (Williams et al., 2013).

Skewness and kurtosis statistics further supported approximate normality. For digital technology self-efficacy, skewness was -0.616 ($SE = 0.223$) and kurtosis was

0.439 (*SE* = 0.442). Trust in automation had skewness of -0.779 (*SE* = 0.223) and kurtosis of -0.112 (*SE* = 0.442). Locus of control exhibited skewness of -0.322 (*SE* = 0.223) and kurtosis of -0.144 (*SE* = 0.442). All values fell within the commonly accepted ±1 range, suggesting no significant skewness or kurtosis issues.

Visual inspections of the histograms, boxplots, and Q-Q plots for each variable were conducted to supplement the normality tests. The histograms showed near-symmetrical distributions for digital technology self-efficacy, trust in automation, and locus of control. Q-Q plots demonstrated reasonably straight lines with minor deviations at the tails, suggesting slight departures from perfect normality but no serious violations. The boxplots indicated no extreme outliers for any variable, with all data points falling within acceptable ranges. Visualizations of the histogram of regression standardized residuals and the normal P-P plot of regression standardized residuals are presented in Appendix A and Appendix B, respectively.

Assumption Testing for Multiple Regression

To ensure that the data met the assumptions for multiple regression analysis, linearity, multicollinearity, homoscedasticity, independence of residuals, and multivariate outliers were assessed. Linearity was examined through inspection of the normal P-P plot of regression standardized residuals and the residuals scatterplot. The P-P plot demonstrated that the residuals closely followed the diagonal line, indicating that the data met the assumption of linearity. Additionally, the scatterplot showed a random, evenly distributed pattern, suggesting no violation of the linearity assumption. The scatterplots are provided in Appendix B and C. Multicollinearity was evaluated using tolerance and

Variance Inflation Factor (VIF) values. Tolerance values for all predictor variables ranged from .501 to .503, and VIF values were all 1.994, falling well within acceptable limits (VIF < 10), indicating no multicollinearity issues. Homoscedasticity was assessed by reviewing the scatterplot of standardized residuals. The assumption of homoscedasticity was also evaluated by visually inspecting the scatterplot of standardized residuals against standardized predicted values (see Appendix C). The scatterplot showed a random distribution of points without a distinct pattern or funnel-shaped spread, indicating that the assumption of homoscedasticity was satisfactorily met. The independence of residuals was verified using the Durbin-Watson statistic, which yielded a value of 2.014. This value is close to the ideal value of 2.0, indicating that residuals were independent, and autocorrelation was not present. Finally, Mahalanobis distance was used to evaluate multivariate outliers. The maximum Mahalanobis distance observed was 14.130, which is below the critical chi-square value of 16.27 for three predictor variables at $p < .001$. No multivariate outliers remained in the dataset after screening.

Reliability of Measurements

Reliability analyses were conducted to assess the internal consistency of the measurement instruments. Cronbach's alpha coefficients were $\alpha = .597$ for the Digital Technology Self-Efficacy Scale, $\alpha = .854$ for the Rotter Locus of Control Scale, and $\alpha = .930$ for the Trust in Automation Scale. These results indicate acceptable to excellent internal consistency across the scales used in this study (see Table 4). Cronbach's alpha for the Digital Technology Self-Efficacy Scale ($\alpha = .597$) reflects the internal consistency of the scale after reverse-coding negatively worded items. While this value is below the commonly accepted threshold of .70, it is considered acceptable for exploratory research (Taber, 2018).

Table 4

Cronbach's Alpha Coefficients for Study Measures (N = 118)

Scale	Number of Items	Cronbach's Alpha
Digital Technology Self-Efficacy Scale	17	.597
Rotter Locus of Control Scale	29	.854
Trust in Automation Scale	12	.930

Multiple Regression Analysis

Standard multiple regression analysis was then conducted to examine whether digital technology self-efficacy, locus of control, and their interaction significantly predicted trust in automation. The predictor variables—digital technology self-efficacy and locus of control—were mean-centered prior to analysis. An interaction term was then created by multiplying the mean-centered predictors, resulting in the self-efficacy x locus of control variable. All three predictors were entered simultaneously into the regression model.

The overall regression model was not statistically significant, $F(3, 114) = 2.451$, $p = .067$, $R^2 = .061$, Adjusted $R^2 = .036$; see Tables 6-8). The main effect of digital technology self-efficacy on trust in automation was not statistically significant, $b = 1.189$, $\beta = .501$, $t(114) = .696$, $p = .488$, suggesting no unique predictive value for digital technology self-efficacy. Similarly, the main effect of locus of control was not statistically significant, $b = 0.671$, $\beta = .253$, $t(114) = .311$, $p = .756$. The interaction term (self-efficacy x locus of control) was also not statistically significant, $b = -0.018$, $\beta = -0.381$, $t(114) = -0.379$, $p = .705$.

Based on these results, I failed to reject the null hypotheses. Neither digital technology self-efficacy, locus of control, nor their interaction significantly predicted trust in automation in this sample. The Durbin-Watson statistic was 2.014, supporting the assumption of independence of residuals. Additionally, inspection of residual scatterplots and normal probability plots indicated the assumptions of homoscedasticity and normality were met.

Table 5

Model Summary for Regression Predicting Trust in Automation

Model	R	R^2	Adjusted R^2	Std. Error of the Estimate	Change in R^2	F Change	$df1$	$df2$	p-value	Durbin-Watson
1	.246	.061	.036	12.70	.061	2.45	3	114	.067	2.01

Note. Predictors: (Constant), Digital Technology Self-Efficacy, Locus of Control, Digital Technology Self-Efficacy × Locus of Control. Dependent Variable: Trust in Automation.

Table 6

ANOVA Summary for Regression Predicting Trust in Automation

Source	Sum of Squares	df	Mean Square	F	p
Regression	1,185.21	3	395.07	2.45	.067
Residual	18,372.73	114	161.16		
Total	19,557.94	117			

Note. Dependent Variable: Trust in Automation. Predictors: (Constant), Digital Technology Self-Efficacy, Locus of Control, Digital Technology Self-Efficacy × Locus of Control.

Table 7

Coefficients and Collinearity Diagnostics for Predictors of Trust in Automation

Variable	*B*	SE *B*	*β*	*t*	*p*
Constant	14.39	76.33	–	0.189	.851
Digital Technology Self-Efficacy	1.19	1.71	.501	0.696	.488
Locus of Control	0.67	2.16	.253	0.311	.756
Self-Efficacy × Locus of Control	-0.02	0.05	-.381	-0.379	.705

Note. Dependent variable: Trust in AI.

Summary

Standard multiple regression analysis was used to examine relationships between digital technology self-efficacy, locus of control, and trust in automation, including the potential moderation effect of locus of control. This indicated digital technology self-efficacy, locus of control, and their interactions did not significantly predict trust in automation. None of the predictors individually or significantly contributed to the model. These findings suggest digital technology self-efficacy and locus of control did not significantly predict trust in automation, either independently or through their interactions. Chapter 5 includes interpretation of findings as well as limitations, implications, and directions for future research and positive social change.

Chapter 5: Discussion, Conclusions, and Recommendations

The purpose of this study was to determine the extent to which locus of control moderated the relationship between digital technology self-efficacy and trust in AI among English-speaking U.S. adults between 22 and 55. With AI increasingly embedded in critical sectors such as healthcare, customer service, and the workplace, understanding psychological underpinnings of user trust has become essential (Esteva et al., 2019; Hancock et al., 2011). Trust significantly influences AI acceptance, particularly in high-stakes contexts where system reliability directly affects outcomes, patient safety and operational success (Guznov et al., 2020; Unver & Asan, 2022). Self-efficacy and locus of control are influential predictors of technology use and acceptance (Bandura, 1977; Rotter, 1966). Individuals with higher self-efficacy typically demonstrate greater confidence and trust when interacting with technology (Hancock et al., 2011; Hsu et al., 2007). Conversely, those with a strong internal locus of control might be more cautious in terms of adopting systems that are perceived as limiting personal autonomy (Chiou et al., 2021). Existing studies have not explored how these traits interact to influence trust in AI.

This quantitative cross-sectional study involved addressing this gap through a survey of 118 participants who were recruited via MTurk. I employed validated measures involving digital technology self-efficacy, locus of control, and trust in AI. Correlation analyses indicated no significant direct relationships between digital technology self-efficacy, locus of control, and trust in AI. A regression model incorporating digital technology self-efficacy, locus of control, and their interactions similarly revealed no significant predictors of trust in AI, nor was there evidence of moderation in terms of locus of control.

Findings from this study highlight the complexity underlying psychological determinants of trust in AI technologies, suggesting predictive relationships involving these factors may be more context-dependent or nuanced than previously understood. In Chapter 5, I interpret these results within the context of existing literature, explore their implications, discuss the study's limitations, and provide recommendations for future research and practice.

Interpretation of the Findings

Digital Technology Self-Efficacy

For RQ1, results indicated digital technology self-efficacy did not significantly predict trust in AI within the regression model that included locus of control and their interactions. Therefore, I failed to reject the null hypothesis. Individuals who feel more competent interacting with technology are typically less anxious and more trusting toward AI systems (Hancock et al., 2011; Hsu et al., 2007). Self-efficacy reduces uncertainty and perceived risks associated with novel technologies, potentially increasing individual openness and willingness to adopt automated systems (Lee & See, 2004; Thatcher et al., 2018).

Absence of significant findings may indicate that influence on trust in AI is contextually contingent on other psychological or situational variables. For instance, locus of control, which is related to beliefs about personal autonomy, could theoretically influence how individuals interpret their technological competence. Chiou et al. (2021)

noted in specific contexts, individuals with an internal locus of control might perceive certain AI systems as potential threats to autonomy. Internal locus of control typically equips individuals with resilience and confidence when interacting with new technologies (Ajzen, 2002; Rotter, 1966). Conversely, individuals with high self-efficacy combined with an external locus of control orientation might more readily trust automation due to greater comfort delegating decisions to external systems (Parasuraman & Riley, 1997).

Additionally, existing literature emphasizes context and specificity of AI applications as critical factors shaping trust dynamics. Trust in AI is not uniform across scenarios and can vary significantly based on perceived task complexity, criticality of systems, and familiarity with specific technological functions. Thus, digital technology self-efficacy may positively influence trust primarily in contexts where individuals perceive alignment between their competencies and technology, therefore feeling genuinely capable of understanding and influencing AI outputs. When systems are perceived as opaque, complex, or high-risk, confidence via self-efficacy alone might diminish and be overridden by perceived system transparency or accountability (Madhavan & Wiegmann, 2007; Lyons & Stokes, 2012).

Locus of Control

In response to RQ2, results indicated no significant direct relationship. Contrary to existing research, participants' locus of control, whether more internal or external, did not significantly predict their level of trust in AI; therefore, I failed to reject the null hypothesis. Individuals with an internal locus of control typically display greater caution or skepticism toward automated technologies due to concerns about losing personal autonomy or control (Chiou et al., 2021; Dzindolet et al., 2003). There may be several explanations which may clarify this divergence. First, locus of control may interact more strongly with contextual or situational variables than previously assumed, making its impact less pronounced in terms of general measures of trust. Trust in AI varies considerably based on specific contexts, including perceived reliability, transparency, and criticality of AI systems (Hoff & Bashir, 2015; Lee & See, 2004). Thus, locus of control might be more predictive in scenarios which explicitly involve personal autonomy or direct consequences to individuals, such as healthcare or financial decision-making, rather than broader or more abstract measures of trust that were used in this study.

Second, contemporary exposure to automation technologies, particularly among individuals who frequently participate in online surveys through platforms such as MTurk, might have moderated the expected effect of locus of control. Frequent and routine interactions with automation could reduce salience of autonomy concerns for individuals with a strong internal locus of control, thereby diminishing observable differences in trust levels between individuals with internal and external locus of control orientations (Glikson & Woolley, 2020). Chiou et al. (2021) suggested extensive exposure to automation technologies might lead individuals with an internal locus of control to recalibrate expectations regarding autonomy. Individuals with an internal locus of control maintain their sense of agency and confidence when navigating new challenges (Ajzen, 2002; Rotter, 1966). Influence of locus of control on technology adoption might be indirect or mediated through perceived usefulness, perceived ease of use, or anxiety (Ajzen, 2002; Venkatesh et al., 2012). Therefore, the absence of a direct relationship in this study does not preclude more complex indirect relationships that alternative analytic approaches such as mediation analyses might be used to address.

Moderating Role of Locus of Control

For RQ3, results revealed no significant moderation effect; therefore, I failed to reject the null hypothesis. Contrary to expectations, locus of control did not alter the strength or direction of the relationship between digital technology self-efficacy and trust in AI in this sample. This finding contrasts with existing literature. Locus of control might influence the degree to which self-efficacy beliefs translate into trust, especially in contexts where personal autonomy and control over outcomes are salient (Chiou et al., 2021; Parasuraman & Riley, 1997).

Several plausible explanations may clarify the absence of this moderation effect. First, the influence of locus of control on technology-related trust may depend heavily upon context-specific factors such as the perceived risk, complexity, or autonomy implications associated with particular AI systems (Hoff & Bashir, 2015). The broad measure of trust in AI used in the current study may have diluted context-specific effects, obscuring subtle moderation by locus of control that might emerge more clearly in narrower or more targeted contexts such as medical decision-making, autonomous driving, or financial advising. Second, the lack of moderation could reflect a more complex relationship involving unmeasured mediating variables. For instance, perceived transparency or explainability of the AI system could mediate or conditionally influence how locus of control impacts trust formation (Shin, 2021; Lyons & Stokes, 2012). Individuals with an internal locus of control, typically valuing autonomy, may depend heavily on system transparency to form trust judgments, whereas externally oriented individuals might be more accepting of opaque automated processes. Without explicitly measuring and incorporating such mediators, these nuanced relationships may remain hidden.

Finally, demographic characteristics of the MTurk sample used in this study (i.e., predominantly younger, well-educated, and highly familiar with digital technologies) could also have influenced moderation outcomes. Regular and extensive experience with digital platforms and AI-driven tools may reduce individual differences related to autonomy or control concerns, thereby minimizing potential moderating effects of locus of control (Glikson & Woolley, 2020; Madhavan & Wiegmann, 2007). This interpretation underscores the importance of investigating psychological variables in conjunction with clearly defined contextual and situational factors.

Limitations of the Study

Several limitations should be considered when interpreting this study's findings. First, participants were recruited via MTurk, resulting in a convenience sample of younger, digitally literate adults, potentially limiting generalizability to broader populations. Research indicates MTurk participants may differ from the general population, often demonstrating higher digital literacy and familiarity with technology-based tasks, which can influence results and limit external validity (Walter et al., 2019). Second, the internal consistency of the Digital Technology Self-Efficacy Scale was lower than the typically recommended threshold. To verify this reliability result, an additional analysis was conducted using raw scores before reverse-coding negatively worded items, which yielded an artificially inflated reliability estimate. This outcome occurred because SPSS treats all items as oriented in the same direction, thus improperly inflating reliability when reverse-coding is not correctly applied. The final reliability therefore accurately reflected the intended scoring procedure of the scale. The relatively low internal consistency likely reflected reduced variance within the sample, as MTurk participants typically possess higher levels of digital self-efficacy compared to general populations, which could attenuate reliability estimates (Chandler & Shapiro, 2016).

Third, reliance on self-report measures and a cross-sectional design introduced additional limitations. Self-report surveys may be susceptible to response biases, particularly social desirability bias, the tendency for respondents to present themselves in a favorable or socially acceptable light (Grimm, 2010). Although participants were assured of anonymity, subjective interpretation and self-presentation concerns may still have influenced responses, affecting accuracy. Finally, the absence of a measurement tool specifically designed to assess trust in AI at the time this study was launched, posed a limitation. The Trust in Automation Scale used in this study was originally developed for general automation systems. To clarify its application within this research, participants were instructed that in the following questions, "automation" referred specifically to AI tools and systems. However, participants' individual interpretations of "AI" may have varied significantly based on their own experience and understanding. Although examples of AI (e.g., voice assistants, chatbots, recommendation systems, productivity tools, autonomous systems) were provided earlier in the survey as part of a demographic question about frequent interactions, these examples were not explicitly reiterated in the trust-related questions. Without employing a measurement tool specifically designed to assess trust in distinct types of AI technologies, or explicitly defining narrower categories within the trust measure itself, this variability in interpretation might have influenced the consistency and accuracy of trust measurements.

Recommendations

Although research on psychological determinants of trust in AI has increased, important areas remain underexplored. The present study investigated the extent to which locus of control moderated the relationship between digital technology self-efficacy and trust in AI. Future research could enhance the robustness and generalizability of these findings by replicating this study with more varied demographic populations beyond the MTurk platform, such as professionals in high-stakes fields (e.g., healthcare, aviation, finance), diverse age groups, or populations with lower digital literacy. Previous studies indicated that demographic diversity significantly impacts trust dynamics and psychological interactions with technology (Glikson & Woolley, 2020; Hoff & Bashir, 2015).

Additionally, future research could incorporate longitudinal designs or analytic approaches such as path analysis to examine how relationships among digital technology self-efficacy, locus of control, and trust in AI evolve over time. Such approaches could provide deeper insight into temporal dynamics and directional influences within these psychological constructs, addressing limitations associated with the analytic strategy employed in this study (Madhavan & Wiegmann, 2007; Parasuraman & Riley, 1997). Further exploration of additional contextual and psychological variables is recommended.

Variables such as perceived transparency, risk perception, system explainability, and specific situational contexts (e.g., healthcare decision-making, autonomous transportation, financial advising) should be examined to provide more granular insights into the trust-automation relationship (Shin, 2021; Lyons & Stokes, 2012).

Given the measurement limitations highlighted, future research should prioritize refining existing scales or developing new validated measures designed for assessing trust specifically in AI technologies. Narrowing the measurement focus to distinct AI types or clearly defined contexts (e.g., specific tools such as chatbots, voice assistants, autonomous vehicles) would enhance consistency and interpretive precision of trust assessments. Employing scales specifically adapted for digital contexts or diverse populations could further improve reliability and validity when studying technologically literate or varied demographic groups.

Future studies should also carefully consider both psychological and contextual factors to fully understand when and why digital technology self-efficacy predicts trust in AI. Investigating the interplay of competence perceptions, autonomy concerns, and contextual specificity can provide more precise guidance on designing trustworthy AI systems tailored to individual psychological profiles. Furthermore, examining context-specific AI applications, exploring indirect pathways linking locus of control to trust, or comparing populations with varying levels of digital exposure would clarify the nuanced roles psychological constructs play in shaping human trust in AI. Employing mediation and moderation analyses with additional variables such as anxiety, perceived usefulness, and task complexity could illuminate indirect pathways and further elucidate psychological mechanisms underlying trust formation (Venkatesh et al., 2012).

Implications

The findings from this study provided meaningful implications for understanding the psychological dynamics involved in trust toward AI technologies. Although the study found no significant direct relationships or moderation effects involving digital technology self-efficacy and locus of control, these results themselves hold important implications for theory and practice. At an individual level, the absence of significant predictive relationships suggested that factors influencing trust in AI might be more nuanced or context-specific than previously assumed. Prior literature frequently positions self-efficacy as a critical factor promoting technology acceptance (Lee & See, 2004; Thatcher et al., 2018); however, this study indicated that general digital self-efficacy might not uniformly translate to increased trust in AI technologies. Thus, it may be important for future research and practical initiatives to explore which specific aspects of technological competence most effectively foster trust in AI, rather than assuming universal applicability across all technology contexts or user groups.

At an organizational level, results implied that generalized measures of psychological traits, such as locus of control, might have limited direct utility in predicting employees' trust in AI across broad settings. This finding aligns with literature emphasizing that trust in automation often depends on situational variables such as system transparency, complexity, perceived risk, or task specificity (Hoff & Bashir, 2015; Madhavan & Wiegmann, 2007). Consequently, organizations, particularly those in critical sectors like healthcare or finance, could prioritize contextual factors, such as the nature of tasks automated or clarity of AI processes, when assessing potential trust barriers among employees, rather than solely focusing on individual psychological traits in isolation.

At a broader theoretical and methodological level, this study's findings highlighted important limitations in existing measurement approaches, particularly regarding trust in automation. Given that participants' individual interpretations of "AI" likely varied widely, the study underscored the necessity of developing more precise and context-specific tools for assessing trust explicitly in AI systems. Such methodological refinements could enhance both research accuracy and the practical relevance of findings, ultimately contributing to clearer and more actionable insights into human-AI interactions.

Conclusion

As artificial intelligence increasingly reshapes critical sectors, including healthcare, transportation, military applications, and finance, understanding the factors that influence user trust has become increasingly important. Automation and AI systems offer transformative benefits, yet widespread adoption continues to face barriers such as misinformation, algorithmic bias, and user anxiety, all of which undermine public confidence. Trust becomes particularly essential in high-stakes environments, where reluctance to rely on automated systems can lead to compromised safety, reduced efficiency, or operational challenges.

This study examined how psychological traits, specifically digital technology self-efficacy and locus of control, relate to trust in AI. Findings indicated that neither digital technology self-efficacy nor locus of control significantly predicted trust in AI, nor did locus of control moderate the relationship between digital technology self-efficacy and trust. Although these results did not confirm initial hypotheses, they highlighted the complexity of psychological influences on trust in automated systems, suggesting that such relationships may be more nuanced, indirect, or context-dependent than previously thought. While the study did not yield direct predictive relationships, the findings still hold theoretical value by challenging existing assumptions about straightforward psychological influences on trust. They emphasized the importance of context specificity and the necessity for future research to explore more nuanced and precise mechanisms underlying human trust in AI technologies.

Ultimately, although this research did not clarify direct dynamics between self-efficacy, locus of control, and trust, it contributed by highlighting important measurement and methodological considerations for future studies. As AI technologies continue to pervade essential sectors of society, this research reinforced the importance of pursuing deeper, context-specific investigations into how psychological traits interact with situational variables to shape trust. Given the transformative potential, and significant risks, of widespread AI adoption, uncovering these nuanced psychological dynamics is not merely valuable; it is imperative to ensure technology genuinely serves and empowers those it aims to assist.

References

Aguinis, H., Beaty, J. C., Boik, R. J., & Pierce, C. A. (2005). Effect size and power in assessing moderating effects of categorical variables using multiple regression: A 30-year review. *Journal of Applied Psychology, 90*(1), 94–107. https://doi.org/10.1037/0021-9010.90.1.94

Ajzen, I. (2002). Perceived behavioral control, self-efficacy, locus of control, and the theory of planned behavior. *Journal of Applied Social Psychology, 32*(4), 665–683. https://doi.org/10.1111/j.1559-1816.2002.tb00236.x

Al-Gasawneh, J. A., Hammouri, Q., Anuar, M. M., & Nusairat, N. M. (2023). Moderating and mediating roles of users' self-efficacy and user behavior amongst the internet of things and search engine optimization. *2023 International Conference on Business Analytics for Technology and Security (ICBATS)*, 1–6. https://doi.org/10.1109/ICBATS57792.2023.10111300

Alkaissi, H., & McFarlane, S. I. (2023). Artificial hallucinations in ChatGPT: Implications in scientific writing. *Cureus*. https://doi.org/10.7759/cureus.35179

Ang, S., Van Dyne, L., & Begley, T. M. (2003). The employment relationships of foreign workers versus local employees: A field study of organizational justice, job satisfaction, performance, and OCB. *Journal of Organizational Behavior*, *24*(5), 561–583. https://doi.org/10.1002/job.202

Bakshy, E., Messing, S., & Adamic, L. A. (2015). Exposure to ideologically diverse news and opinion on Facebook. *Science*, *348*(6239), 1130–1132. https://doi.org/10.1126/science.aaa1160

Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review*, *84*(2), 191–215. https://doi.org/10.1037/0033-295X.84.2.191

Bandura, A. (1982). Self-efficacy mechanism in human agency. *American Psychologist, 37*(2), 122–147. https://doi.org/10.1037/0003-066X.37.2.122

Bandura, A., & National Inst of Mental Health. (1986). *Social foundations of thought and action: A social cognitive theory.* Prentice-Hall.

Bandura, A. (1989). Social cognitive theory. In R. Vasta (Ed.), *Six theories of child development* (Vol. 6, pp. 1-60). JAI Press.

Bandura, A. (1997). *Self-efficacy: The exercise of control.* W H Freeman.

Bandura, A. (2012). On the functional properties of perceived self-efficacy revisited [Editorial]. *Journal of Management, 38*(1), 9–44. https://doi.org/10.1177/0149206311410606

Bandura, A. (2006). Guide for constructing self-efficacy scales. In F. Pajares & T. Urdan (Eds.), *Self-efficacy and adolescence* (pp. 307–337). Information Age Publishing.

Bandura, A., & Walters, R. H. (1963). *Social learning and personality development.* Holt, Rinehart, & Winston.

Barocas, S., & Selbst, A. (2016). *Big data's disparate impact*. https://doi.org/10.15779/Z38BG31

Beck, A. T. (1961). An inventory for measuring depression. *Archives of General Psychiatry*, *4*(6), 561. https://doi.org/10.1001/archpsyc.1961.01710120031004

Beretvas, S. N., Suizzo, M. A., Durham, J. A., & Yarnell, L. M. (2008). A reliability generalization study of scores on Rotter's and Nowicki-Strickland's Locus of Control Scales. *Educational and Psychological Measurement*, *68*(1), 97-119. https://doi.org/10.1177/0013164407301529

Berry, J. W., Poortinga, Y. H., Segall, M. H., & Dasen, P. R. (1992). *Cross-cultural psychology: Research and applications.* Cambridge University Press.

Bickmore, T. W., & Picard, R. W. (2005). Establishing and maintaining long-term human-computer relationships. *ACM Transactions on Computer-Human Interaction*, *12*(2), 293–327. https://doi.org/10.1145/1067860.1067867

Bolukbasi, T., Chang, K.-W., Zou, J. Y., Saligrama, V., & Kalai, A. T. (2016). Man is to computer programmer as woman is to homemaker? Debiasing word embeddings. In D. Lee, M. Sugiyama, U. Luxburg, I. Guyon, & R. Garnett (Eds.), *Advances in neural information processing systems* (Vol. 29). https://proceedings.neurips.cc/paper_files/paper/2016/file/a486cd07e4ac3d270571622f4f316ec5-Paper.pdf

Buhrmester, M., Kwang, T., & Gosling, S. D. (2016). Amazon's Mechanical Turk: A new source of inexpensive, yet high-quality data? In A. E. Kazdin (Ed.), *Methodological issues and strategies in clinical research* (4th ed., pp. 133-139). American Psychological Association. https://doi.org/10.1037/14805-009

Buolamwini, J., & Gebru, T. (2018). Gender shades: Intersectional accuracy disparities in commercial gender classification. *Proceedings of Machine Learning Research, 81*, 77–91.

Burkhardt, M. E., & Brass, D. J. (1990). Changing patterns or patterns of change: The effects of a change in technology on social network structure and power. *Administrative Science Quarterly*, *35*(1), 104. https://doi.org/10.2307/2393552

Cassidy, S., & Eachus, P. (2002). Developing the Computer User Self-Efficacy (CUSE) scale: Investigating the relationship between computer self-efficacy, gender, and experience with computers. *Journal of Educational Computing Research, 26*(2), 133–153. https://doi.org/10.2190/JGJR-0KVL-HRF7-GCNV

Chandler, J., & Shapiro, D. (2016). Conducting clinical research using crowdsourced convenience samples. *Annual Review of Clinical Psychology, 12*(1), 53–81. https://doi.org/10.1146/annurev-clinpsy-021815-093623

Chaturvedi, R., & Verma, S. (2023). Artificial intelligence-driven customer experience: Overcoming the challenges. *California Management Review. 65(*3), 77–91.

Chen, T. D., Kockelman, K. M., & Hanna, J. P. (2016). Operations of a shared, autonomous, electric vehicle fleet: Implications of vehicle & charging infrastructure decisions. *Transportation Research Part A: Policy and Practice*, *94*, 243–254. https://doi.org/10.1016/j.tra.2016.08.020

Cheng, R., Wang, R., Zimmermann, T., & Ford, D. (2022). *It would work for me too: How online communities shape software developers' trust in AI-powered code generation tools*. https://doi.org/10.48550/ARXIV.2212.03491

Chesney, B., & Citron, D. (2019). Deep fakes: a looming challenge for privacy. *California Law Review Online, 108*, 25–32. https://doi.org/10.15779/Z38RV0D15J

Chien, S.-Y., Sycara, K., Liu, J.-S., & Kumru, A. (2016). Relation between trust attitudes toward automation, Hofstede's cultural dimensions, and big five personality traits. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *60*(1), 841–845. https://doi.org/10.1177/1541931213601192

Chiou, M., McCabe, F., Grigoriou, M., & Stolkin, R. (2021). Trust, shared understanding and locus of control in mixed-initiative robotic systems. *2021 30th IEEE International Conference on Robot & Human Interactive Communication (RO-MAN)*, 684–691. https://doi.org/10.1109/RO-MAN50785.2021.9515476

Choi, Y. (2021). A study of employee acceptance of artificial intelligence technology. *European Journal of Management and Business Economics*, *30*(3), 318–330. https://doi.org/10.1108/EJMBE-06-2020-0158

Chou, W.-Y. S., Oh, A., & Klein, W. M. P. (2018). Addressing health-related misinformation on social media. *JAMA*, *320*(23), 2417. https://doi.org/10.1001/jama.2018.16865

Ciechanowski, L., Przegalinska, A., Magnuski, M., & Gloor, P. (2019). In the shades of the uncanny valley: An experimental study of human–chatbot interaction. *Future Generation Computer Systems*, *92*, 539–548. https://doi.org/10.1016/j.future.2018.01.055

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Erlbaum.

Compeau, D. R., & Higgins, C. A. (1995). Computer self-efficacy: Development of a measure and initial test. *MIS Quarterly*, *19*(2), 189. https://doi.org/10.2307/249688

Creswell, J.W., & Creswell, J.D. (2018). Research design: Qualitative, quantitative, and mixed methods approaches (5th ed.). SAGE.

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

Dostovalov, S. G. (2013). Trust in the social world as a factor in the formation of legal consciousness in adolescents. *Psychology and Law, 3*(1), 1-11. https://psyjournals.ru/journals/psylaw/archive/2013_n1/58316

Dzindolet, M. T., Peterson, S. A., Pomranky, R. A., Pierce, L. G., & Beck, H. P. (2003). The role of trust in AI reliance. *International Journal of Human-Computer Studies, 58*(6), 697–718. https://doi.org/10.1016/S1071-5819(03)00038-7

Eastwood, B. (2024). The who, what, and where of AI adoption in America. *MIT Sloan School of Management.* https://mitsloan.mit.edu/ideas-made-to-matter/who-what-and-where-ai-adoption-america

Edwards, C., Edwards, A., Spence, P. R., & Shelton, A. K. (2014). Is that a bot running the social media feed? Testing the differences in perceptions of communication quality for a human agent and a bot agent on Twitter. *Computers in Human Behavior*, *33*, 372–376. https://doi.org/10.1016/j.chb.2013.08.013

Esteva, A., Robicquet, A., Ramsundar, B., Kuleshov, V., DePristo, M., Chou, K., Cui, C., Corrado, G., Thrun, S., & Dean, J. (2019). A guide to deep learning in healthcare. *Nature Medicine*, *25*(1), 24–29. https://doi.org/10.1038/s41591-018-0316-z

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2009). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods, 39*(2), 175-191. https://doi.org/10.3758/BF03193146

Ferring, D., & Filipp, S.-H. (1996). Messung des Selbstwertgefühls: Befunde zu Reliabilität, Validität und Stabilität der Rosenberg-Skala. *Diagnostica, 42*, 284–292.

Floridi, L. (2019). *The ethics of artificial intelligence: Principles, challenges, and opportunities*. Oxford University Press.

Furnham, A., & Steele, H. (1993). Measuring locus of control: A critique of general, children's health- and work-related locus of control questionnaires. *British Journal of Psychology, 84*(4), 443–479. https://doi.org/10.1111/j.2044-8295.1993.tb02495.x

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

Gratch, J., Lucas, G. M., King, A. A., & Morency, L. P. (2014). It's only a computer: The impact of human-agent interaction in clinical interviews. *Proceedings of the 2014 International Conference on Autonomous Agents and Multi-Agent Systems*, 85–92.

Grimm, P. (2010). Social desirability bias. *Wiley international encyclopedia of marketing*. https://doi.org/10.1002/9781444316568.wiem02057

Groh, M., Epstein, Z., Obradovich, N., Cebrian, M., & Rahwan, I. (2021). Human detection of machine-manipulated media. *Communications of the ACM*, *64*(10), 40–47. https://doi.org/10.1145/3445972

Grove, R. M., Fowler, F. J., Jr., Couper, M. P., Lepkowski, J. M., Singer, E., &

Tourangeau, R. (2009). *Survey Methodology* (2nd ed.). Wiley.

Guznov, S., Lyons, J., Pfahler, M., Heironimus, A., Woolley, M., Friedman, J., & Neimeier, A. (2020). Robot transparency and team orientation effects on human–robot teaming. *International Journal of Human–Computer Interaction*, *36*(7), 650–660. https://doi.org/10.1080/10447318.2019.1676519

Hancock, P. A., Billings, D. R., Schaefer, K. E., Chen, J. Y. C., De Visser, E. J., & Parasuraman, R. (2011). A meta-analysis of factors affecting trust in human-robot interaction. *Human Factors*, *53*(5), 517–527. https://doi.org/10.1177/0018720811417254

Hatamlah, H. (2024). Adolescents' online health information seeking: Trust, e-health literacy, parental influence, and AI-generated credibility. *International Journal of Data and Network Science*, *8*(2), 809–822. https://doi.org/10.5267/j.ijdns.2023.12.023

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

Hofstede, G. (1991). *Cultures and organizations: Software of the Mind*. McGraw-Hill.

Holcomb, L., King, F., & Brown, S. (2004). Student traits and attributes contributing to success in online courses: evaluation of university online courses. *Journal of Interactive Online Learning*, *2*.

Howard, A., & Borenstein, J. (2018). The ugly truth about ourselves and our robot creations: The problem of bias and social inequity. *Science and Engineering*

Ethics, *24*(5), 1521–1536. https://doi.org/10.1007/s11948-017-9975-2

Hughes, J. E. (2013). Descriptive indicators of future teachers' technology integration in the PK-12 classroom: Trends from a laptop-infused teacher education program. *Journal of Educational Computing Research, 48*(4), 493–518. https://doi.org/10.2190/ec.48.4.e

Hurt, H. T., Joseph, K., & Cook, C. D. (1977). Scales for the measurement of innovativeness. *Human Communication Research*, *4*(1), 58–65.

https://doi.org/10.1111/j.1468-2958.1977.tb00597.x

Hsia, J. W., Chang, C. C., & Tseng, A. H. (2014). Effects of individuals' locus of control and computer self-efficacy on their e-learning acceptance in high-tech companies. *Behaviour & Information Technology*, *33*(1), 51-64.

Hsu, M.-H., Ju, T. L., Yen, C.-H., & Chang, C.-M. (2007). Knowledge sharing behavior in virtual communities: The relationship between trust, self-efficacy, and outcome expectations. *International Journal of Human-Computer Studies*, *65*(2), 153–169. https://doi.org/10.1016/j.ijhcs.2006.09.003

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

Jerusalem, M., & Schwarzer, R. (1992). Self-efficacy as a resource factor in stress appraisal processes. In R. Schwarzer (Ed.), *Self-efficacy: Thought control of action* (pp. 195–213). Hemisphere Publishing Corp.

Jian, J.-Y., Bisantz, A. M., & Drury, C. G. (2000). Foundations for an empirically determined scale of trust in automated systems. *International Journal of Cognitive Ergonomics*, *4*(1), 53–71. https://doi.org/10.1207/S15327566IJCE0401_04

John, O. P., & Srivastava, S. (1999). The big five trait taxonomy: history, measurement, and theoretical perspectives. In *Handbook of personality: Theory and research, (*2nd ed., pp. 102–138). Guilford Press.

Katz, J. E., & Halpern, D. (2014). Attitudes towards robots suitability for various jobs as affected robot appearance. *Behaviour & Information Technology*, *33*(9), 941–953. https://doi.org/10.1080/0144929X.2013.783115

Körber, M. (2018). Theoretical considerations and development of a questionnaire to measure trust in automation. In S. Bagnara, R. Tartaglia, S. Albolino, T. Alexander, & Y. Fujita (Eds.), *Proceedings of the 20th Congress of the International Ergonomics Association (IEA 2018): Volume VI: Transport Ergonomics and Human Factors (TEHF), Aerospace Human Factors and Ergonomics* (pp. 13–30). Springer.

Kraus, J., Scholz, D., Messner, E.-M., Messner, M., & Baumann, M. (2020). Scared to trust? – predicting trust in highly automated driving by depressiveness, negative self-evaluations and state anxiety. *Frontiers in Psychology*, *10*, 2917. https://doi.org/10.3389/fpsyg.2019.02917

Kuznetsova, E., Makhortykh, M., Vziatysheva, V., Stolze, M., Baghumyan, A., & Urman, A. (2023). *In generative ai we trust: can chatbots effectively verify political information?* https://doi.org/10.48550/ARXIV.2312.13096

Kreps, S., McCain, R. M., & Brundage, M. (2022). All the news that's fit to fabricate: ai-generated text as a tool of media misinformation. *Journal of Experimental Political Science*, *9*(1), 104–117. https://doi.org/10.1017/XPS.2020.37

Lazer, D. M. J., Baum, M. A., Benkler, Y., Berinsky, A. J., Greenhill, K. M., Menczer, F., Metzger, M. J., Nyhan, B., Pennycook, G., Rothschild, D., Schudson, M., Sloman, S. A., Sunstein, C. R., Thorson, E. A., Watts, D. J., & Zittrain, J. L. (2018). The science of fake news. *Science*, *359*(6380), 1094–1096. https://doi.org/10.1126/science.aao2998

Lee, J. D., & See, K. A. (2004). Trust in automation: Designing for appropriate reliance. *Human Factors*, *46*(1), 50–80. https://doi.org/10.1518/hfes.46.1.50_30392

Lefcourt, H. M., & Davidson-Katz, K. (1991). Locus of control and health. In C. R. Snyder & D. R. Forsyth (Eds.), *Handbook of social and clinical psychology: The health perspective* (pp. 246–266). Pergamon Press.

Leontiev, A. A., & Leontiev, D. A. (Eds.). (1994). *Filosofiia psikhologii: iz nauchnogo naslediia*. Moscow Smysl; IP RAE

Lestari, N. S., Rosman, D., Chan, S., Nawangsari, L. C., Natalina, H. D., & Triono, F. (2022). Impact of robots, artificial intelligence, service automation (raisa) acceptance, self-efficacy, and relationship quality on job performance. *2022 4th International Conference on Cybernetics and Intelligent System (ICORIS)*, 1–6. https://doi.org/10.1109/ICORIS56080.2022.10031336

Liu, T., Zhang, Y., Brockett, C., Mao, Y., Sui, Z., Chen, W., & Dolan, B. (2021, April). A token-level reference-free hallucination detection benchmark for free-form text generation. *ACL2022 Main Conference*.

Lucas, G. M., Gratch, J., King, A., & Morency, L.-P. (2014). It's only a computer: Virtual humans increase willingness to disclose. *Computers in Human Behavior*, *37*, 94–100. https://doi.org/10.1016/j.chb.2014.04.043

Luger, E., & Sellen, A. (2016). "Like having a really bad PA": the gulf between user expectation and experience of conversational agents. *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, 5286–5297. https://doi.org/10.1145/2858036.2858288

Lyons, J. B., & Stokes, C. K. (2012). Human–human reliance in the context of automation. *Human Factors, 54*(1), 112–121. https://doi.org/10.1177/0018720811427034

Maslow, A. H. (1954). *Motivation and personality*. Harper & Row.

Madhavan, P., & Wiegmann, D. A. (2007). Similarities and differences between human-human and human-automation trust: An integrative review. *Theoretical Issues in Ergonomics Science, 8*(4), 277–301. https://doi.org/10.1080/14639220500337708

Marteau, T. M., & Bekker, H. (1992). The development of a six-item short-form of the state scale of the Spielberger State—Trait Anxiety Inventory (STAI). *British Journal of Clinical Psychology*, *31*(3), 301–306.
https://doi.org/10.1111/j.2044-8260.1992.tb00997.x

Marakas, G. M., Yi, M. Y., & Johnson, R. D. (1998). The multilevel and multifaceted character of computer self-efficacy: Toward clarification of the construct and an integrative framework for Research. *Information Systems Research*, *9*(2), 126–

163. https://doi.org/10.1287/isre.9.2.126

Merritt, S. M., Heimbaugh, H., LaChapell, J., & Lee, D. (2013). I trust it, but I don't know why: effects of implicit attitudes toward automation on trust in an automated system. *Human Factors*, *55*(3), 520–534. https://doi.org/10.1177/0018720812465081

Miller, G. A., Galanter, E., & Pribram, K. H. (1960). *Plans and the structure of behavior.* Henry Holt and Co. https://doi.org/10.1037/10039-000

Mittelstadt, B. D., Allo, P., Taddeo, M., Wachter, S., & Floridi, L. (2016). The ethics of algorithms: Mapping the debate. *Big Data & Society, 3*(2). https://doi.org/10.1177/2053951716679679

Merritt, S. M., & Ilgen, D. R. (2008). Not all trust is created equal: Dispositional and history-based trust in human-automation interactions. *Human Factors*, *50*(2), 194–210. https://doi.org/10.1518/001872008X288574

Mlekus, L., Bentler, D., Paruzel, A., Kato-Beiderwieden, A.-L., & Maier, G. W. (2020). How to raise technology acceptance: User experience characteristics as technology-inherent determinants. *Gruppe. Interaktion. Organisation. Zeitschrift Für Angewandte Organisationspsychologie (GIO)*, *51*(3), 273–283. https://doi.org/10.1007/s11612-020-00529-7

Montag, C., Kraus, J., Baumann, M., & Rozgonjuk, D. (2023). The propensity to trust in (automated) technology mediates the links between technology self-efficacy and fear and acceptance of artificial intelligence. *Computers in Human Behavior Reports*, *11*, 100315. https://doi.org/10.1016/j.chbr.2023.100315

Murphy, C. A., Coover, D., & Owen, S. V. (1989). Development and validation of the computer self-efficacy scale. *Educational and Psychological Measurement*, *49*(4), 893–899. https://doi.org/10.1177/001316448904900412

Neisser, U. (2014). *Cognitive psychology: Classic edition*. Psychology press.

Norman, C. D., & Skinner, H. A. (2006). eHEALS: The eHealth Literacy Scale. *Journal of Medical Internet Research*, *8*(4), e27. https://doi.org/10.2196/jmir.8.4.e27

Novozhilova, E., Mays, K., & Katz, J. E. (2024). Looking towards an automated future: U.S. attitudes towards future artificial intelligence instantiations and their effect. *Humanities and Social Sciences Communications*, *11*(1), 132. https://doi.org/10.1057/s41599-024-02625-1

O'Donovan, N. (2020). From knowledge economy to automation anxiety: A growth regime in crisis? *New Political Economy*, *25*(2), 248–266. https://doi.org/10.1080/13563467.2019.1590326

Ok, M. W., Hughes, J. E., & Lee, B.-I. (2017). A study of effects of 1:1 laptop computing on pre-service special educators' technology-related knowledge, skills, and attitudes. *The Korean Journal of Early Childhood Special Education, 17*(2), 1–21. http://dx.doi.org/10.21214/kecse.2017.17.2.1

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*, 568256. https://doi.org/10.3389/fpsyg.2020.568256

Olmstead, K. (2017) Nearly half of Americans use digital voice assistants, mostly on their smartphones. *Pew Research Center*. https://www.pewresearch.org/short-reads/2017/12/12/nearly-half-of-americans-use-digital-voice-assistants-mostly-on-their-smartphones/

Opdahl, A. L., Tessem, B., Dang-Nguyen, D.-T., Motta, E., Setty, V., Throndsen, E., Tverberg, A., & Trattner, C. (2023). Trustworthy journalism through AI. *Data & Knowledge Engineering*, *146*, 102182. https://doi.org/10.1016/j.datak.2023.102182

Pajares, F. (1996). Self-efficacy beliefs and mathematical problem-solving of gifted students. *Contemporary Educational Psychology*, *21*(4), 325–344. https://doi.org/10.1006/ceps.1996.0025

Pajares, F., & Miller, M. D. (1994). Role of self-efficacy and self-concept beliefs in mathematical problem solving: A path analysis. *Journal of Educational Psychology, 86*(2), 193–203. https://doi.org/10.1037/0022-0663.86.2.193

Parasuraman, R., & Riley, V. (1997). Humans and automation: Use, misuse, disuse, abuse. *Human Factors, 39*(2), 230–253. https://doi.org/10.1518/001872097778543886

Pennycook, G., & Rand, D. G. (2019). Fighting misinformation on social media using crowdsourced judgments of news source quality. *Proceedings of the National Academy of Sciences*, *116*(7), 2521–2526. https://doi.org/10.1073/pnas.1806781116

Piercy, C., & Gist-Mackey, A. (2021). Automation anxieties: Perceptions about technological automation and the future of pharmacy work. *Human-Machine*

Communication, *2*, 191–208. https://doi.org/10.30658/hmc.2.10

Rahwan, I., Cebrian, M., Obradovich, N., Bongard, J., Bonnefon, J.-F., Breazeal, C., Crandall, J. W., Christakis, N. A., Couzin, I. D., Jackson, M. O., Jennings, N. R., Kamar, E., Kloumann, I. M., Larochelle, H., Lazer, D., McElreath, R., Mislove, A., Parkes, D. C., Pentland, A. 'Sandy,' … Wellman, M. (2019). Machine behaviour. *Nature*, *568*(7753), 477–486. https://doi.org/10.1038/s41586-019-1138-y

Riedl, R. (2022). Is trust in artificial intelligence systems related to user personality? Review of empirical evidence and future research directions. *Electronic Markets*, *32*(4), 2021–2051. https://doi.org/10.1007/s12525-022-00594-4

Rizzo, J. R., House, R. J., & Lirtzman, S. I. (1970). Role conflict and ambiguity in complex organizations. *Administrative Science Quarterly*, *15*(2), 150. https://doi.org/10.2307/2391486

Rogers, C. R. (1951). *Client-centered therapy; its current practice, implications, and theory.* Houghton Mifflin.

Rotter, J. B. (1966). Generalized expectancies for internal versus external control of reinforcement. *Psychological Monographs: General and Applied*, *80*(1), 1–28. https://doi.org/10.1037/h0092976

Rotter, J. B. (1971). Generalized expectancies for interpersonal trust. *American Psychologist*, *26*(5), 443–452. https://doi.org/10.1037/h0031464

Rowley, J., Johnson, F., & Sbaffi, L. (2015). Students' trust judgements in online health information seeking. *Health Informatics Journal*, *21*(4), 316–327.

https://doi.org/10.1177/1460458214546772

Salvagno, M., Taccone, F. S., & Gerli, A. G. (2023). Artificial intelligence hallucinations. *Critical Care*, *27*(1), 180. https://doi.org/10.1186/s13054-023-04473-y

Scherer, R., Siddiq, F., & Teo, T. (2015). Becoming more specific: Measuring and modeling teachers' perceived usefulness of ICT in the context of teaching and learning. *Computers & Education*, *88*, 202–214. https://doi.org/10.1016/j.compedu.2015.05.005

Schunk, D. H. (1991). Self-efficacy and academic motivation. *Educational Psychologist, 26*(3-4), 207–231. https://doi.org/10.1207/s15326985ep2603&4_2

Schyns, B., & Von Collani, G. (2002). A new occupational self-efficacy scale and its relation to personality constructs and organizational variables. *European Journal of Work and Organizational Psychology*, *11*(2), 219-241.

Selezneva, Y., Abakumova, I., & Kupriyanov, I. (2023). On the question of a human's personality resources in a changing world: volitional control, trust, anxiety. *International Journal of Cognitive Research in Science, Engineering and Education*, *11*(2), 291–300. https://doi.org/10.23947/2334-8496-2023-11-2-291-300

Shao, C., Ciampaglia, G. L., Varol, O., Yang, K.-C., Flammini, A., & Menczer, F. (2018). The spread of low-credibility content by social bots. *Nature Communications*, *9*(1), 4787. https://doi.org/10.1038/s41467-018-06930-7

Shank, D. B., DeSanti, A., & Maninger, T. (2019). When are artificial intelligence versus human agents faulted for wrongdoing? Moral attributions after individual and joint decisions. *Information, Communication & Society*, *22*(5), 648–663. https://doi.org/10.1080/1369118X.2019.1568515

Shapkin, S. A. (1997). *Experimental study of volitional processes*. Moscow: Smysl; IP RAE.

Sharan, N. N., & Romano, D. M. (2020). The effects of personality and locus of control on trust in humans versus artificial intelligence. *Heliyon*, *6*(8), e04572. https://doi.org/10.1016/j.heliyon.2020.e04572

Shin, D. (2021). The effects of explainability and causability on perception, trust, and acceptance: Implications for explainable AI. *International Journal of Human-Computer Studies, 146*, 102551. http://dx.doi.org/10.1016/j.ijhcs.2020.102551

Shin, D., Zaid, B., Biocca, F., & Rasul, A. (2022). In platforms we trust? Unlocking the black-box of news algorithms through interpretable AI. *Journal of Broadcasting & Electronic Media*, *66*(2), 235–256. https://doi.org/10.1080/08838151.2022.2057984

Sindermann, C., Sha, P., Zhou, M., Wernicke, J., Schmitt, H. S., Li, M., Sariyska, R., Stavrou, M., Becker, B., & Montag, C. (2021). Assessing the attitude towards artificial intelligence: Introduction of a short measure in German, Chinese, and English language. *KI - Künstliche Intelligenz*, *35*(1), 109–118. https://doi.org/10.1007/s13218-020-00689-0

Skripkina, T. P., & Selezneva, Y. V. (2014). Deformation of trusting relationships as a factor in the development of professional deformations of a teacher of preschool educational organizations. *Theoretical and Experimental Psychology, 7*(1), 24-29.

Stajkovic, A. D., & Luthans, F. (1998). Self-efficacy and work-related performance: A meta-analysis. *Psychological Bulletin, 124*(2), 240–261.

https://doi.org/10.1037/0033-2909.124.2.240

Taber, K. S. (2018). The use of Cronbach's alpha when developing and reporting research instruments in science education. *Research in Science Education, 48*(6), 1273–1296. https://doi.org/10.1007/s11165-016-9602-2

Taffesse, A. S., & Tadesse, F. (2017). Pathways less explored—Locus of control and technology adoption. *Journal of African Economies*, *26*(suppl_1), i36-i72.

Thatcher, J. B., Zimmer, J. C., Gundlach, M. J., & McKnight, D. H. (2018). Internal and external dimensions of computer self-efficacy: An empirical examination. *IEEE Transactions on Engineering Management, 65*(1), 33-45. https://doi.org/10.1109/TEM.2008.927825

Thompson, R. L., Higgins, C. A., & Howell, J. M. (1991). Personal computing: Toward a conceptual model of utilization. *MIS Quarterly*, *15*(1), 125. https://doi.org/10.2307/249443

To, Q. G., Green, C., & Vandelanotte, C. (2021). Feasibility, usability, and effectiveness of a machine learning–based physical activity chatbot: Quasi-experimental study. *JMIR mHealth and uHealth*, *9*(11), e28577. https://doi.org/10.2196/28577

Torkzadeh, G., Koufteros, X., & Pflughoeft, K. (2003). Confirmatory analysis of computer self-efficacy. *Structural Equation Modeling: A Multidisciplinary Journal*, *10*(2), 263–275. https://doi.org/10.1207/S15328007SEM1002_6

Unver, B.M., & Asan, O. (2022). Role of trust in AI-driven healthcare systems:

Discussion from the perspective of patient safety. *Proceedings of the International Symposium on Human Factors and Ergonomics in Health Care*, *11*(1), 129–134. https://doi.org/10.1177/2327857922111026

U.S. Census Bureau (2023). *Age and sex composition in the United States: 2023*. https://www.census.gov/data/tables/2023/demo/age-and-sex/2023-age-sex-composition.html

U.S. Census Bureau (2021). Language spoken at home: 2021 American Community Survey. https://www.census.gov/topics/population/language-use.html

Van Dijk, J. A. G. M. (2017). Digital divide: impact of access. In P. Rössler, C. A. Hoffner, & L. Zoonen (Eds.), *The International Encyclopedia of Media Effects* (1st ed., pp. 1–11). Wiley. https://doi.org/10.1002/9781118783764.wbieme0043

Venkatesh, V., & Davis, F. (2000). A theoretical extension of the technology acceptance model: Four longitudinal field studies. *Management Science*, *46*, 186–204.

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

Vieira, L. N. (2020). Automation anxiety and translators. *Translation Studies*, *13*(1), 1–21. https://doi.org/10.1080/14781700.2018.1543613

Walter, S. L., Seibert, S. E., Goering, D., & O'Boyle, E. H., Jr. (2019). A tale of two sample sources: Do results from online panel data and conventional data converge? *Journal of Business and Psychology, 34*(4), 425–452. https://doi.org/10.1007/s10869-018-9552-y

Warner, R. M. (2013). *Applied statistics: From bivariate through multivariate techniques* (2nd ed.). Sage.

Williams, M. N., Grajales, C. A. G., & Kurkiewicz, D. (2013). Assumptions of multiple regression: Correcting two misconceptions. *Practical Assessment, Research, and Evaluation*, *18*(1), 11.

Williams, P., Soutar, G., Ashill, N. J., & Naumann, E. (2017). Value drivers and adventure tourism: A comparative analysis of Japanese and Western consumers. *Journal of Service Theory and Practice*, *27*(1), 102–122. https://doi.org/10.1108/JSTP-05-2015-0116

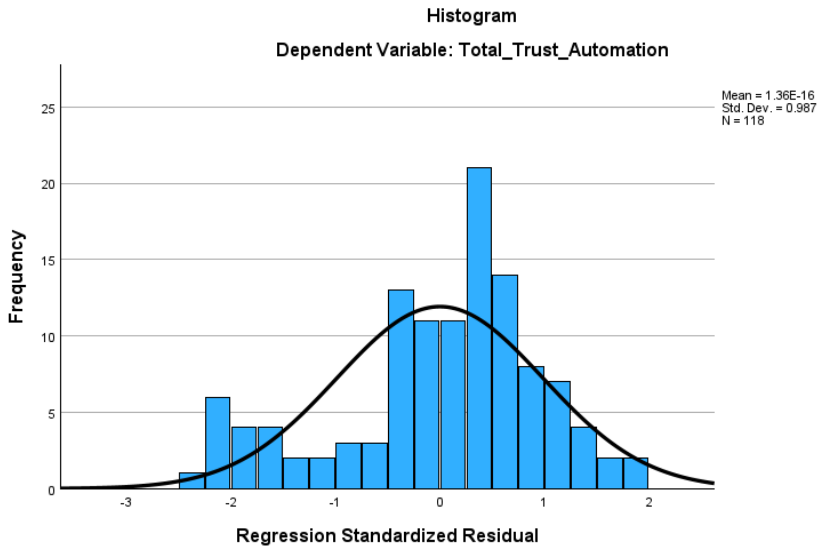
Zajko, M. (2021). Conservative AI and social inequality: Conceptualizing alternatives to bias through social theory. *AI & Society*, *36*(3), 1047–1056. https://doi.org/10.1007/s00146-021-01153-9

Zimmerman, B. J., Bandura, A., & Martinez-Pons, M. (1992). Self-motivation for academic attainment: the role of self-efficacy beliefs and personal goal setting. *American Educational Research Journal*, *29*(3), 663–676. https://doi.org/10.3102/00028312029003663

Zimmerman, B. J. (2000). Self-efficacy: An essential motive to learn. *Contemporary Educational Psychology*, *25*(1), 82–91. https://doi.org/10.1006/ceps.1999.1016

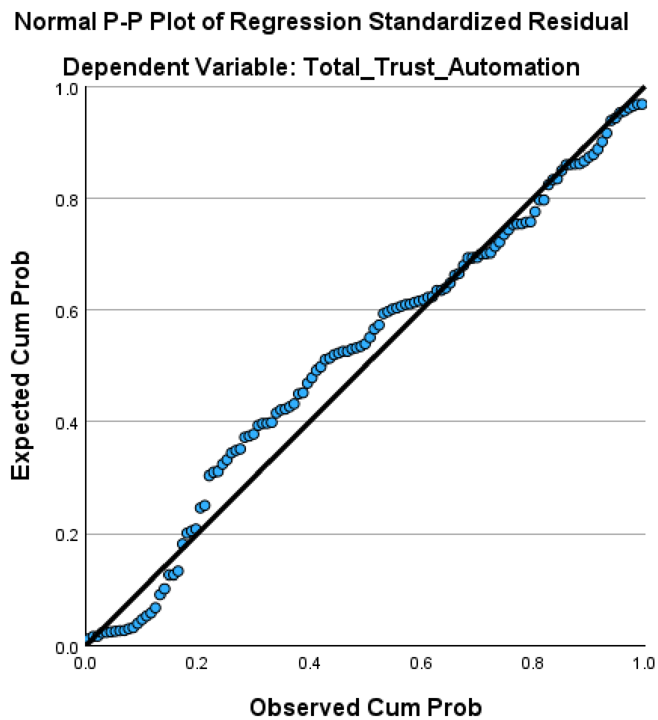
Appendix A

Histogram of Regression Standardized Residuals



Appendix B

Normal P-P Plot of Regression Standardized Residuals



Appendix C

Scatterplot of Regression Standardized Residuals and Predicted Values