



Is trust in artificial intelligence systems related to user personality? Review of empirical evidence and future research directions

René Riedl^{1,2}

Received: 11 February 2022 / Accepted: 7 September 2022
© The Author(s) 2022, corrected publication 2023

Abstract

Artificial intelligence (AI) refers to technologies which support the execution of tasks normally requiring human intelligence (e.g., visual perception, speech recognition, or decision-making). Examples for AI systems are chatbots, robots, or autonomous vehicles, all of which have become an important phenomenon in the economy and society. Determining which AI system to trust and which not to trust is critical, because such systems carry out tasks autonomously and influence human decision making. This growing importance of trust in AI systems has paralleled another trend: the increasing understanding that user personality is related to trust, thereby affecting the acceptance and adoption of AI systems. We developed a framework of user personality and trust in AI systems which distinguishes universal personality traits (e.g., Big Five), specific personality traits (e.g., propensity to trust), general behavioral tendencies (e.g., trust in a specific AI system), and specific behaviors (e.g., adherence to the recommendation of an AI system in a decision-making context). Based on this framework, we reviewed the scientific literature. We analyzed $N=58$ empirical studies published in various scientific disciplines and developed a “big picture” view, revealing significant relationships between personality traits and trust in AI systems. However, our review also shows several unexplored research areas. In particular, it was found that prescriptive knowledge about how to design trustworthy AI systems as a function of user personality lags far behind descriptive knowledge about the use and trust effects of AI systems. Based on these findings, we discuss possible directions for future research, including adaptive systems as focus of future design science research.

Keywords Artificial Intelligence (AI) · Big Five traits · Machine learning (ML) · Personality · Review · Trust · Trust propensity

JEL classification D91 · M15 · O33

Responsible Editor: Wolfgang Maass

✉ René Riedl
rene.riedl@fh-steyr.at; rene.riedl@jku.at

¹ Digital Business Institute, School of Business and Management, University of Applied Sciences Upper Austria, Wehrgrabengasse 1-3, 4400 Steyr, Austria

² Institute of Business Informatics – Information Engineering, Johannes Kepler University Linz, Altenberger Strasse 69, 4040 Linz, Austria

Introduction

McCarthy et al. (1955) referred to Artificial Intelligence (AI) as “making a machine behave in ways that would be called intelligent if a human were so behaving” (p. 11). In his seminal book on AI history, Nilsson (2010) specifies that intelligence is “that quality that enables an entity to function appropriately and with foresight in its environment [... and] because ‘functioning appropriately and with foresight’ requires so many different capabilities, depending on the environment, we actually have several continua of intelligences” (book preface). Thus, despite the fact that we observe an ongoing discussion on AI definitions in the scientific literature (that mainly results from the way how intelligence is conceptualized), in recent years AI systems such as chatbots, recommendation agents, autonomous vehicles,

and robots have become a dominating topic, both in science and practice (e.g., Berente et al., 2021; Dwivedi et al., 2021). The success of integrating AI systems into society and the economy significantly depends on users' trust (e.g., Glikson & Woolley, 2020; Jacovi et al., 2021). AI systems increasingly carry out tasks autonomously and significantly influence human-decision making. Hence, determining which AI system to trust and which not to trust has become crucial. Trust in systems and machines which turn out not to be trustworthy may lead to undesired consequences (Lee & Moray, 1992). These may range from unfavorable purchases (e.g., when using recommendation agents in online shops) to death (e.g., when an autonomous vehicle causes a fatal accident). What follows is that the decision to trust AI systems always comes with some form of vulnerability. This explains why many novel AI systems are not always embraced by people, despite their enormous potential to make organizations more efficient and human life more comfortable (Collins et al., 2021). Considering this *enormous relevance of users' trust for the acceptance and adoption of AI systems*, it is no surprise that scientific research has studied this phenomenon extensively. In the past decade, several reviews on trust in autonomous systems and AI systems were published (Glikson & Woolley, 2020; Hoff & Bashir, 2015; Siau & Wang, 2018; Thiebes et al., 2021), and also meta-analyses are available (Hancock et al., 2011; Schaefer et al., 2016).

This growing importance of trust in AI systems has paralleled another trend: the *increasing relevance of personality research in the fields of AI, robotics, and autonomous systems* (Matthews et al., 2021). The American Psychological Association (APA) defines personality as "individual differences in characteristic patterns of thinking, feeling and behaving".¹ Other conceptualizations highlight the stability of the characteristics over time. For example, Mount et al. (2005) indicate that personality traits can be defined as "characteristics that are *stable over time* [...] which] provide the reasons for the person's behavior [...] they reflect who we are and in aggregate determine our affective, behavioral, and cognitive style" (pp. 448-449, italics added). In a more recent paper, Montag and Panksepp (2017) define personality as a set of "*stable individual differences in cognitive, emotional and motivational aspects of mental states that result in stable behavioral action* [...] especially emotional tendencies" (p. 1, italics added).

Considering these definitions, it is obvious that human personality and trust tendencies must be related, a fact that also holds true in the specific context of trust in AI systems. However, to date, the role of personality in the context of trust in AI systems has *not* been investigated in the form of

a systematic review of existing empirical evidence. Rather, existing works only briefly mention user personality as a factor related to trust in automation and AI systems (Glikson & Woolley, 2020; Hoff & Bashir, 2015; Siau & Wang, 2018; Thiebes et al., 2021). While this confirms the relevance of this article's topic, these brief references to user personality do not constitute systematic analysis of the existing knowledge. Moreover, recently Matthews et al. (2021) made the following statement in their paper on personality research in robotics, autonomous systems, and AI: "Person factors in trust in machines have been a relatively neglected aspect of research [...]" (p. 3). User personality is a major person factor. In another recent paper, Jacovi et al. (2021) make an explicit call for personality research in the context of trust in AI systems—they write: "Personal Attributes of the Trustor [...] Future work in this area may incorporate elements of the personal attributes of the trustor into the model, such as personality" (p. 633).

The fact that personality research in the context of trust in AI systems has not been examined systematically in the form of a review is problematic for at least four reasons.

First, as it has turned out after analyzing the extant literature, the existing empirical studies are published in various scientific disciplines, including computer science and robotics, ergonomics, human-computer interaction, information systems (IS), and psychology. Hence, a cumulative research tradition hardly exists. What follows is that to date a "*big-picture*" view of user personality and trust in the context of AI systems is not available. Thus, it is a fruitful scientific endeavor to *integrate empirical results across discipline boundaries and contexts*.

Second, investigation of the development of research at the nexus of user personality and trust in AI systems contributes to identity development in this important interdisciplinary and relatively new field. Based on a review of N = 58 empirical papers published in the period 2008-2021, in the current article we integrate available, but highly fragmented literature that is published in outlets across various scientific disciplines. Examination of a research field's development may provide valuable *insights into the future development*. In an essay on the identity of the IS discipline, Klein and Hirschheim (2008) write that "a shared sense of history provides the ultimate grounding and background information (preunderstanding) for communication in large and diverse collectives such as societies (and by extension to diverse disciplines)" (p. 298). Bearing this argument in mind, a main motivation of the present paper is to provide insights into the status of personality research in the context of trust in AI systems, thereby contributing to identity development. Knowing the history of the research field, even if it is short, facilitates identity formation by *identification of emergent thematic and methodological patterns*. In short, knowing

¹ <https://www.apa.org/topics/personality> (accessed on January 16, 2022).

one's past is important for coping with future challenges (Webster & Watson, 2002).

Third, personality has been identified as a major individual difference variable (along with gender, age, and culture) which affects trust in the context of automation and AI systems, and it has been argued that this fact has far-reaching *design consequences* (e.g., Hoff & Bashir, 2015). For example, Tapus et al.'s (2008) robot study demonstrated the feasibility of “a behavior adaptation system capable of adjusting its social interaction parameters (e.g., interaction distances/proxemics, speed, and vocal content) [...] based on the user's personality traits and task performance” (p. 169). Importantly, a *sound understanding of users' personality and trust constitutes a foundation for the successful development of adaptive systems*. In general, such systems automatically adapt in real time based on users' states or traits to improve human-technology interactions (e.g., vom Brocke et al., 2020). Development of adaptive systems has become an important topic in IS research (e.g., Astor et al., 2013; Demazure et al., 2021) and also beyond IS, as signified by developments in disciplines like affective computing (e.g., Poria et al., 2017).

Fourth, a major motivation to study personality research in the context of trust in AI systems is that IS research has ignored this personality focus thus far. Maier (2012) and recently Sindermann et al. (2020) reviewed personality research in the IS discipline and report two thematic domains in which personality has played a major role: as antecedent in technology acceptance and adoption decisions, and as characteristic of computer personnel which influences further downstream variables (e.g., willingness to change job). Regarding acceptance and adoption, it is critical to emphasize that the extant IS literature had a focus on traditional applications, including enterprise systems and online shops without any AI component (e.g., Devaraj et al., 2008; McElroy et al., 2007). Therefore, while personality research has its place in the IS discipline, the present paper's specific thematic focus has not been addressed thus far.

The article is structured as follows. In Sect. 2, we develop a theoretical foundation for the present article. Specifically, we summarize foundations of personality research, as well as major insights into trust in the context of AI systems. In Sect. 3, we outline our review methodology. Sect. 4 presents major results of the review. Sect. 5 discusses the results, identifies several unexplored research areas, and outlines several fruitful directions for future research. In Sect. 6, we specifically discuss design implications and describe the development of adaptive systems as possible focus of future design science research. In Sect. 7, this paper's limitations and a concluding statement are provided. Altogether, the present article contributes to a better understanding of an increasingly important phenomenon in the digital economy and society, namely the influence of user personality on

trust in AI systems, which, in turn, significantly affects the acceptance and adoption decisions. Therefore, the current topic is not only essential from an academic perspective, but also from a practice point of view.

Theoretical foundations

Personality

Human personality refers to stable characteristics that determine individual differences in thoughts, emotions, and behavior (Funder, 2001). Thus, personality affects how people react to all forms of stimuli, including situations in which humans interact with AI systems.

Today consensus exists that the *Five-factor Model*—in short, *Big Five*—constitutes the “broadest level of abstraction” to conceptualize human personality, where “each dimension summarizes a large number of distinct, more specific personality characteristics” (John & Srivastava, 1999, p. 105).² Initially described in the 1960s (Tupes & Christal, 1961), several research groups have contributed independently to the development of the model in the past decades (Cattell et al., 1970; Goldberg, 1990; McCrae & Costa, 1987). The model comprises the following five factors: extraversion (“an energetic approach toward the social and material world”, example traits: sociability, activity, assertiveness, positive emotionality), agreeableness (“contrasts a prosocial and communal orientation toward others with antagonism”, example traits: altruism, tender-mindedness, trust), conscientiousness (“describes socially prescribed impulse control that facilitates task- and goal-directed behavior”, example traits: thinking before acting, delaying gratification, following norms and rules, planning, organizing, and prioritizing tasks), neuroticism³ (“contrasts emotional stability and even-temperedness with negative emotionality”, trait examples: feeling anxious, nervous, sad, tense), and openness⁴ (“describes the breadth, depth, originality, and complexity of an individual's mental and experiential life”, example traits: being imaginative, original, insightful, curious) (conceptualizations and examples taken from John et al., 2008, p. 120; for further details, see McCrae & Costa, 1997, 1999).

² The “Big Five” are a conceptualization which is independent from language and culture; moreover, its structure emerges across observers (self-reports, peer-reports) and methodologies (questionnaires, lexical inventories) (e.g., John & Srivastava, 1999; McCrae, 2004; Mooradian et al., 2006).

³ Emotional instability is used as a synonym in the scientific literature.

⁴ Intellect is used as a synonym in the scientific literature.

In addition to the Big Five, further well-known personality models exist. Based on work by Carl Jung (1923), Myers and Briggs developed a personality model referred to as *Myers-Briggs Type Indicator (MBTI)*. In essence, this model dichotomizes four dimensions: extraversion vs. introversion, sensation vs. intuition, thinking vs. feeling, judging vs. perceiving (Myers et al., 1998). Hence, this conceptualization comes along with $2^4 = 16$ unique personality types. Another framework is the *Eysenck personality model* which is based on two dimensions: extraversion (E) and neuroticism (N) (Eysenck, 1947). Based on high and low levels of each dimension, four personality types emerge: $E_H + N_H$: choleric type, $E_L + N_H$: melancholic type, $E_H + N_L$: sanguine type, $E_L + N_L$: phlegmatic type. Later, psychoticism was added as a third dimension, which can be divided into impulsivity and sensation-seeking (Eysenck & Eysenck, 1976). The *HEXACO model of personality* was developed by Ashton et al. (2004) and is based on the work of Costa Jr. and McCrae (1992) and Goldberg (1993). Hence, this model resembles the Big Five framework. It has six dimensions: honesty-humility, emotionality, extraversion, agreeableness, conscientiousness, and openness to experience. Thus, if compared to the Big Five, HEXACO has a sixth dimension, namely honesty-humility (with the sub-dimensions sincerity, fairness, greed avoidance, and modesty).⁵ Finally, the *Revised NEO Personality Inventory (NEO PI-R)* is a model to conceptualize personality based on the Big Five. However, additionally this inventory includes six sub-factors for each trait. It follows that this model uses 30 factors to conceptualize personality. The most recent version of this inventory is referred to as NEO PI-3 (McCrae et al., 2005) and it includes the following factors: extraversion (warmth, gregariousness, assertiveness, activity, excitement seeking, positive emotions), agreeableness (trust, straightforwardness, altruism, compliance, modesty, tender-mindedness), conscientiousness (competence, order, dutifulness, achievement striving, self-discipline, deliberation), neuroticism (anxiety, hostility, depression, self-consciousness, impulsiveness, vulnerability), and openness to experience (fantasy, aesthetics, feelings, actions, ideas, values).

Trust

In situations of interpersonal interaction, trust is defined as “the willingness of a party to be vulnerable to the actions of another party based on the expectation that the other will perform a particular action important to the trustor, irrespective of the ability to monitor or control that other party” (Mayer et al., 1995, p. 712). Moreover, it is an established

fact that trustworthiness beliefs are significantly influenced by perceptions about the trustee’s ability, benevolence, and integrity (e.g., Mayer et al., 1995; Rousseau et al., 1998). However, trust in machines, computers, and AI systems differs from interpersonal trust. In a recent paper, Riedl (2021) analyzed work by the research groups of David Gefen (e.g., Paravastu et al., 2014) and Harrison McKnight (e.g., Lankton et al., 2015), two well-known IS scholars with a particular focus on trust research, and concludes that more appropriate trusting beliefs in human-technology relationships are performance and functionality (instead of ability), helpfulness (instead of benevolence), and predictability and reliability (instead of integrity).⁶ This conclusion is consistent with works in (e.g., Siau & Wang, 2018; Söllner et al., 2012; Thiebes et al., 2021) and outside (e.g., Hancock et al., 2011; Hoff & Bashir, 2015; Glikson & Woolley, 2020) the IS discipline.⁷

Disposition to trust (synonyms: trust propensity, interpersonal propensity to trust) also plays a major role in explaining behavioral intentions and actual behavior. McKnight et al. (1998) define disposition to trust as “a tendency to be willing to depend on others [...] across a broad spectrum of situations and persons” (p. 474, 477). Disposition to trust is typically conceptualized as an antecedent of more situation-specific trust (e.g., Hoff & Bashir, 2015; McKnight et al., 1998). Thus, whether a user develops trusting beliefs and trusting intentions in, and shows actual trusting behavior toward, an IT artifact (e.g., AI system) is influenced by trust disposition. McKnight et al. (1998) further decompose disposition to trust into faith in humanity (defined as “one believes that others are typically well-meaning and reliable”, p. 477) and trusting stance (defined as “one believes that, regardless of whether people are reliable or not, one will obtain better interpersonal outcomes by dealing with people as though they are well-meaning and reliable”, p. 477). Complementing

⁶ Based on Riedl (2021, Table 3.I.), we define as follows (in verbatim): Performance is the belief about the capability of the technology to accomplish its designated purpose. Functionality is the belief that the technology has the capability, functions, or features to fulfil the requirements. Helpfulness is the belief that the specific technology provides adequate and responsive help for users. Predictability is the belief that the technology will do what it is claimed to do without adding anything malicious on top of it. Reliability is the belief that the technology will consistently operate properly.

⁷ Note that trusting beliefs in the context of automation technology and autonomous systems may also refer to purpose (defined as “the degree to which the automation is being used within the realm of the designer’s intent”) and process (defined as “the degree to which the automation’s algorithms are appropriate for the situation and able to achieve the operator’s goals” (Lee & See, 2004, p. 59). For further details, please see a recent paper by Thiebes et al. (2021) who mapped trusting beliefs (performance, etc.) and AI frameworks (e.g., OECD principles on AI) to five trustworthy AI principles (beneficence, non-maleficence, autonomy, justice, and explicability).

⁵ It is important to note that while the other five dimensions resemble the Big Five, they are not perfect matches.

McKnight et al.'s (1998) work, Gefen (2000) argues that trust disposition “is not based upon experience with or knowledge of a specific trusted party [...] but is the result of an ongoing lifelong experience [...] and socialization [...] disposition to trust is most effective in the initiation phases of a relationship when the parties are still mostly unfamiliar with each other [...] and before extensive ongoing relationships provide a necessary background for the formation of other trust-building beliefs, such as integrity, benevolence, and ability” (p. 728). Consistent with these views, Mayer et al. (1995) argued that “[p]ropensity to trust is proposed to be a stable within-party factor that will affect the likelihood the party will trust [...] People with different developmental experiences, personality types, and cultural backgrounds vary in their propensity to trust” (p. 715).

Framework of personality and trust in AI systems

A fundamental question is how personality is related to trust, and we focus on this important question in the context of AI systems. Moreover, it is essential how both personality and trust affect acceptance and adoption.

Propensity to trust is considered to be a specific trait of agreeableness (e.g., John & Srivastava, 1999; John et al., 2008; McCrae & Costa, 1997, 1999).⁸ What follows is that the more specific personality trait propensity to trust can be deduced from the more abstract trait agreeableness. This fact is consistent with discourse in the literature which indicates that personality traits can be conceptualized on different abstraction levels, both in psychology (e.g., McCrae et al., 2005) and IS (e.g., Maier, 2012). Regarding abstraction levels of personality, John and Srivastava (1999) wrote with reference to the Big Five that “these five dimensions represent personality at the broadest level of abstraction, and each dimension summarizes a large number of distinct, more specific personality characteristics” (p. 105). Models like those from Myers-Briggs, Eysenck, or HEXACO conceptualize personality on a similarly high abstraction level. However, this advantage of broad categories also has a disadvantage, the “low fidelity” (John & Srivastava, 1999, p. 124). Decades ago it has already been argued in the personality literature that choice of abstraction levels depends on the descriptive and predictive tasks to be addressed (Hampson et al., 1987), and “[i]n principle, the number of specific distinctions one can make in the description of an individual is infinite, limited only by one’s objectives” (John & Srivastava, 1999, p. 124).

⁸ Mooradian et al. (2006, p. 527) argue that in personality frameworks such as Costa and McCrae’s NEO framework, authors often use the term “trust”, despite the fact that what is actually meant is “propensity to trust”.

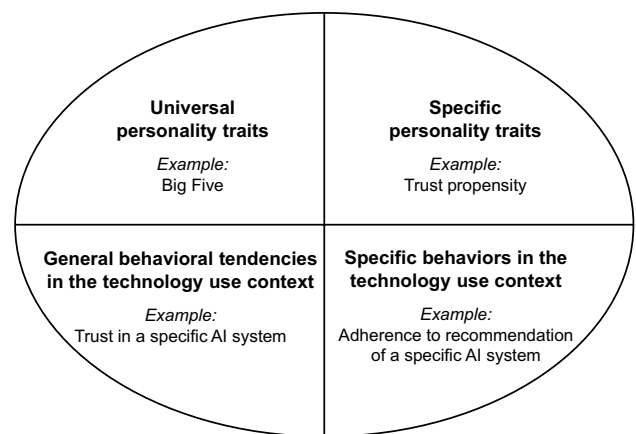


Fig. 1 Framework of user personality and trust in AI systems

Against this background, we devise a framework of personality and trust in AI systems with different abstraction levels. In personality research, it is well-established to include—in addition to the broad category on the highest abstraction level—at least one further level with more specific personality traits, which, in turn, manifest in general behavioral tendencies and more specific behaviors (e.g., Funder, 2001; John & Srivastava, 1999; McCrae et al., 2005).

Figure 1 shows a framework of personality and trust in AI systems. This framework draws upon the logic of existing conceptualizations and models at the nexus of personality and behavior (Funder, 2001; John & Srivastava, 1999; McCrae et al., 2005; Mooradian et al., 2006). In essence, the framework specifies universal personality traits (e.g., Big Five), specific personality traits (e.g., interpersonal propensity to trust), general behavioral tendencies (e.g., trust in a specific AI system such as autonomous vehicles or speech assistants like Alexa or Siri), and specific behaviors (e.g., adherence to the recommendation of an AI system in a decision-making context).

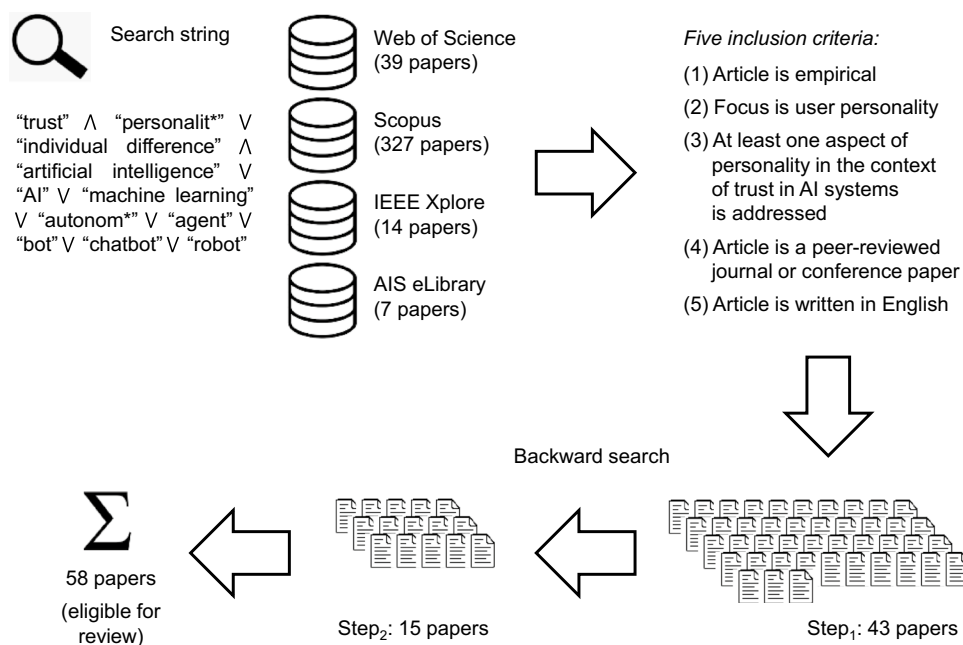
Altogether, our framework draws upon the empirically-grounded facts that (i) personality is related to behavioral tendencies and specific behaviors and (ii) universal personality traits are related to specific personality traits (e.g., Zimbardo et al., 2021). We will refer to the framework in Fig. 1 in the following sections. First, we use the distinction between universal and specific personality traits to code the scientific literature. Second, our discussion of possible future research domains is also based on this framework.

Methodology of the literature review

Literature search

In order to identify publications at the nexus of user personality and trust in the context of AI systems, we conducted a

Fig. 2 Overview of the literature search process



literature search. The search process was based on existing recommendations, in particular vom Brocke et al. (2009). The search was conducted via *Web of Science*, *Scopus*, *IEEE Xplore*, and *AIS eLibrary* starting on 11/28/2021 and a last query was made on 12/31/2021. No publication year restriction was used for all searches.

Step₁: Because AI systems may have different manifestations, we used several keywords. Specifically, we used the following keywords: “trust”, “personalit*”, “individual difference”, “artificial intelligence”, “AI”, “machine learning”, “autonom*”, “agent”, “bot”, “chatbot”, and “robot” (*Web of Science*, *Scopus*, *IEEE Xplore*). For the search in the *AIS eLibrary*, we used “trust”, “personalit*”, and “individual difference”.⁹

This search method resulted in the following number of hits: *Web of Science* = 39 papers, *Scopus* = 327,¹⁰ *IEEE Xplore*: 14, *AIS eLibrary*: 7. Next, we removed duplicates (as the four databases are not mutually exclusive) and read all abstracts of the remaining articles, if necessary also the full text, to identify papers eligible for our review. We applied the following inclusion criteria: (1) the article is empirical

(i.e., it presents collection, analysis, and interpretation of data), (2) the focus is on user personality, and not on the personality of the AI system,¹¹ (3) the article deals with at least one aspect of human personality (e.g., extraversion) in the context of trust in AI systems, (4) the paper constitutes a peer-reviewed journal or conference publication, and (5) the article is written in English. After applying these inclusion criteria, 43 papers remained. Next, we read the full texts of all 43 papers and finally decided to consider them all in our review.

Step₂: Based on the 43 papers, we also conducted a backward search. The same five inclusion criteria were applied as in Step₁. Based on this procedure, we identified an additional 15 papers, all of which were ultimately considered eligible for our review.¹² Hence, the total number of papers included in our review is $N = 58$.¹³ Fig. 2 graphically summarizes the literature search process. Appendix 1 lists the 58 references.

⁹ Note that specifications such as title, topic, abstract, and keywords were not used in exactly the same way across the four databases as the search functions are not identical. Moreover, because the *AIS eLibrary* covers much fewer publication outlets than the other three databases, we did not limit our search to the context of AI in the beginning in order to organize the search more comprehensively.

¹⁰ Note that among the high number of 327 papers identified via *Scopus*, many papers were *not* eligible for consideration in our review as this database also covers conference abstracts (which often describe research-in-progress and only have around 150 words). As a sample abstract, see Panganiban et al. (2020).

¹¹ For example, Hess et al. (2009) investigated how recommendation agent personality (i.e., extraversion) serves as a social technology cue to affect the perceived social presence of the agent. Thus, the focus of this study is not on user personality, but on the personality of the AI system. Hence, this paper just as many other papers with a similar focus on AI system personality were *not* considered in the present review.

¹² Note that we identified one high-quality paper published on *arXiv.org*, namely Sarkar et al. (2017). This platform is a well-known free distribution service and an open-access archive. We decided to consider this paper in our review.

¹³ A forward search did not yield further relevant papers. This is plausible because half of the identified papers (i.e., 29 out of 58) were published in 2020 or 2021, and the searches itself already covered four databases, of which three are among the worldwide largest ones in the domain of the current paper (i.e., *Web of Science*, *Scopus*, *IEEE Xplore*).

Literature coding

To systematically document the existing knowledge on the relationship between user personality and trust in the context of AI systems, as well as meta-information (e.g., publication year), we extracted the following data:

- (1) *General information about the paper*: (a) authors, (b) title, (c) publication year, (d) type of outlet (journal, conference), and (e) scientific discipline, in which a paper appeared. Due to the interdisciplinary nature of the current topic, we defined the following discipline categories ex-ante: (i) Computer Science, Informatics, Robotics, (ii) Ergonomics, Human Factors, (iii) Human-Computer Interaction, (iv) IS, (v) Psychology, and (vi) other.
- (2) *Investigated system*: Because AI systems manifest in various forms, we used different categories, namely: (a) robot (including robot simulations), (b) autonomous vehicle (including in-vehicle agents), (c) user interface (e.g., chatbot), (d) automated agents in a security context (e.g., airport), (e) AI-driven healthcare systems, (f) autonomous system (general), and (g) other (e.g., voice assistant, intelligent tutoring). The categories were developed inductively based on the coded material.
- (3) *Research method*: We grouped the empirical examinations into one of the following categories: (a) lab experiment, (b) survey study (typically conducted online), (c) online experiment, (d) lab study (in contrast to the lab experiment, this type of study does not manipulate an independent variable to systematically study the resulting effects on a dependent variable, but simply observes use of a technology in a laboratory environment). The categories were developed inductively based on the coded material.
- (4) *Sample characteristics*: We documented (a) sample size, (b) mean age of the sample, (c) a sample's gender distribution, and (d) country of data collection.
- (5) *Personality traits*: We documented all investigated personality traits. We distinguished between (a) universal and (b) specific traits. In the category (a), we included (i) papers which investigated all Big Five traits together and (ii) one specific Big Five trait. Moreover, we captured other broad personality conceptualizations, namely (iii) Myers-Briggs type indicator, (iv) Eysenck's model, (v) HEXACO, and (vi) NEO-PI-3. We defined these universal categories ex-ante. In the category (b), we documented all specific traits (e.g., trust propensity) which we found in the analyzed literature.
- (6) *Theoretical framework*: We documented whether a paper included a theoretical framework in the form of

a graphical representation (i.e., constructs and relationships). While the minimum requirement for a theory to exist is at least one independent and one dependent variable, our data analysis revealed that most papers in which a framework was available used more sophisticated theorizing. Specifically, we documented whether the investigated personality construct(s), as well as the trust construct, were conceptualized as independent, mediator, moderator, or dependent variable. Moreover, we documented all dependent variable(s) of a framework (e.g., adoption intention of an AI system).

- (7) *Relationship of personality traits with trust in AI system*: We documented statistically significant relationships between universal traits, as well as specific traits, and trust in AI system. We emphasize that these relationships are typically those which the authors of a paper explicitly highlighted in the abstract and/or the discussion section of a paper. Based on this procedure, we guarantee a focus on the most important results from the perspective of the papers' authors.

Results

We structure the presentation of our results into two subsections. We start with the descriptive results (i.e., point (1) to (4) as listed in the previous literature coding section), followed by our findings on the investigated personality traits and underlying theoretical models, point (5) to (7) in the list.

Descriptive results

Appendix 1 lists the 58 references. Appendix 2 summarizes the major characteristics of these 58 papers and their underlying studies. The 58 papers report 64 studies. Based on the information in Appendix 1 and 2, Table 1 summarizes descriptive statistics of our review. Specifically, we report publication year, the type of outlet in which the papers are published, and the scientific discipline in which a paper appeared.¹⁴ Moreover, we document the investigated technology, research method, and country of data collection. For further sample characteristics (i.e., sample size, mean age of the sample, and rate of females in the sample) we refer the reader to Appendix 2.

¹⁴ Coding of the 58 papers regarding scientific discipline was a straightforward process. However, with respect to the Hawaii International Conference on System Sciences (HICSS) we decided to assign this publication venue to the IS discipline (and not to Computer Science, Informatics, and Robotics). The reason for this decision is that that this conference focuses on "Information Technology *Management*" (see <https://aisel.aisnet.org/hicss/>, italics added).

Table 1 Descriptive statistics

Metric	Findings													
<i>Publication year</i> (N = 58 papers)	08	09	10	11	12	13	14	15	16	17	18	19	20	21
	2	0	0	2	2	2	0	5	2	4	4	6	17	12
<i>Publication outlet</i> (N = 58 papers)	Conference proceedings													30
	Journal													27
	arXiv													1
<i>Scientific Discipline</i> (N = 58 papers)	Computer Science, Informatics, Robotics													18
	Ergonomics, Human Factors													13
	Human-Computer Interaction													11
	Information Systems (IS)													8
	Psychology													4
	Other (e.g., Communication, Sociology)													4
<i>Investigated system</i> (N = 59 studies for which the system is reported)	Robot (including robot simulations)													20
	Autonomous vehicle (including in-vehicle agents)													12
	User interface (e.g., chatbot)													11
	Automated agent in security context (e.g., airport)													7
	AI-driven healthcare systems													3
	Autonomous system and AI (general)													3
	Other (e.g., voice assistant, intelligent tutoring)													3
<i>Research method</i> (N = 60 studies for which the method is reported)	Laboratory experiment													22
	Survey study													18
	Online experiment													14
	Laboratory study													6
<i>Country (data collection)</i> (N = 60 studies for which the country is reported)	USA													30
	Mix of different countries													8
	Germany													6
	United Kingdom													4
	Canada													3
	Italy													2
	Australia													2
	Japan													2
	China													1
	Korea													1
	South Africa													1

Findings on personality traits

Table 2 summarizes our findings on the investigated personality traits, separated into universal and specific traits.

Regarding the universal traits, we observed that 18 out of the 58 papers studied the Big Five. Moreover, in 11 further papers, single Big Five traits were studied. While extraversion was studied in 5 papers, neuroticism was studied in 3, agreeableness in 2, and openness in 1 paper, respectively. No article had a focus on conscientiousness alone. Moreover, the results show that other general personality models only play a minor role in the extant literature. Specifically, the Myers-Briggs and Eysenck models, as well as HEXACO and NEO-PI-3, have only been applied once each.

Table 2 further indicates that we identified 33 specific personality traits.¹⁵ We list all traits which were studied in at least two papers (i.e., 12 traits).¹⁶ We found that trust propensity (i.e., the disposition to trust other people) and dispositional trust in automation/machines are the most studied

¹⁵ Note that we harmonized terminology across the 58 papers in our analyses and tables. For example, we use the term “trust propensity” despite the fact that synonyms like “interpersonal propensity to trust”, “disposition to trust”, “dispositional trust”, “dispositional interpersonal trust”, “dispositional trust in humans”, or “trust disposition” are used in the literature.

¹⁶ In Appendix 2 (see column “Specific traits”) the reader can find the 21 traits which were only examined in one of the 58 papers.

Table 2 Statistics on the personality traits

Traits	Findings	
<i>Universal personality traits</i> (N = 58 papers)	Big Five (together)	18
	Extraversion	5
	Neuroticism	3
	Agreeableness	2
	Openness	1
	Conscientiousness	0
	Myers-Briggs	1
	Eysenck	1
	HEXACO	1
	NEO-PI-3	1
	<i>Specific personality traits</i> (N = 58 papers)	Trust propensity
Dispositional trust in automation/machines		9
Locus of control		7
Attitude towards computers/robots		5
Self-efficacy		5
Technological innovativeness		4
Need for cognition		3
Perfect automation expectation		3
Propensity to take risks		3
Sensation seeking		3
Need for interaction		2
Self-esteem		2
Other (e.g., computer anxiety, thrill seeking)		21

specific personality traits (in 9 papers each). Also, locus of control plays a significant role in the literature (studied in 7 papers). Locus of control is the degree to which people believe that they have control over the outcome of events (as opposed to external forces beyond their influence); this locus can either be internal (a person believes that he or she can control their own life) or external (the belief that factors which cannot be influenced control life) (Rotter, 1966, 1990). Both attitude towards computers/robots and self-efficacy (the degree to which people believe that they can accomplish a particular task or activity, Bandura, 1977, Gist, 1987) were studied in 5 papers each. The following traits were examined in three papers each: need for cognition (“an individual’s tendency to engage in and enjoy effortful cognitive endeavors”, Cacioppo et al., 1984, p. 306), perfect automation expectation (“expectations that the automated aid [e.g., autonomous vehicle] will perform with near-perfect reliability”, Merritt et al., 2015, p. 740),¹⁷ propensity to take risks, and sensation seeking (“the tendency to seek novel, varied, complex, and intense sensations and experiences and the willingness to take risks for the sake of such experiences”, Choi & Ji, 2015, p. 694). Finally, need for

interaction and self-esteem were studied in 2 papers each. Need for interaction is a person’s preference to stay in touch with other people during the use of a service or application (Dabholkar, 1992). Self-esteem is defined as “the level of global regard one has for the self as a person” (Harter, 1993, p. 88).

Regarding the availability of theoretical frameworks, we found graphical representations with constructs and relationships in 19 out of the 58 papers.¹⁸ Table 3 summarizes the results of our analyses. As shown in Table 3, we found that 18 out of the 19 papers conceptualize personality traits as independent variable. Another major result is that trust is conceptualized as independent variable in 3 papers only. However, trust is much more often conceptualized either as mediator (7 times) or as dependent variable (11 times). Table 3 further shows that personality traits are rarely conceptualized as mediator (2 times) or moderator (3 times). Also, we observe that in studies in which trust is not the dependent variable, often behavioral intention to use an AI system is used as outcome variable.

Regarding the relationship between personality traits and trust in AI systems, a first notable result was that many of the hypothesized relationships in the analyzed papers were statistically *not* significant. This concerns both the

¹⁷ Note that this expectation is related to the concept of perfect automation schema (PAS), which has two dimensions: first, high expectations for automation performance, and second, the fact that users hardly forgive computers and machines when they make mistakes. For further details, see Dzindolet et al. (2002).

¹⁸ We only focused on frameworks with direct relevance for the present article (i.e., the framework has at least one personality construct and one trust construct).

Table 3 Analysis of the theoretical conceptualization of personality and trust (N=19 papers in which graphical representations of theoretical frameworks are provided)

Reference	Pers=IV	Trust=IV	Pers=MED	Trust=MED	Pers=MOD	Trust=MOD	Pers=DV	Trust=DV	DV
Bawack et al. (2021)	x			x					Customer experience performance
Choi and Ji (2015)	x	x							Behavioral intention
Cohen & Sergay (2011)	x	x							Behavioral intention
Elson et al. (2018)				x	x				Decision
Elson et al. (2020)	x							x	
Handrich (2021)	x			x					Adoption intention
Hanna & Richards (2015)	x							x	
Hegner et al. (2019)	x	x		x					Adoption intention
Kim et al. (2020)	x							x	
Kraus et al. (2020a)	x		x					x	
Kraus et al. (2020b)	x							x	
Matthews et al. (2020)	x							x	
Merritt et al. (2013)	x				x			x	
Miller et al. (2021)	x			x				x	
Schaefer & Scribner (2015)	x			x					Performance
Sharan & Romano (2020)	x							x	
Youn & Jin (2021)	x		x		x			x	Intention to visit, satisfaction
Zhang et al. (2020)	x			x					Behavioral intention
Zhou et al. (2020)	x							x	
Sum	18	3	2	7	3	0	0	11	

IV independent variable, MED mediator variable, MOD moderator variable, DV dependent variable, Pers Personality

Big Five traits and the more specific traits. However, we also observed relationships which turned out to be stable across the analyzed papers.

With respect to the *Big Five*, we found the following results:

- Agreeableness: 9 papers found a positive relationship with trust in AI systems (Bawack et al. 2021, Böckle et al. 2021, Chien et al. 2016, Ferronato & Bashir 2020a, Huang et al. 2020, Kraus et al. 2020a, Lyons et al. 2020, Müller et al. 2019, Rossi et al. 2018).
- Openness: 8 papers found a positive relationship with trust in AI systems (Aliasghari et al. 2021, Antes et al. 2021, Böckle et al. 2021, Elson et al. 2020, Ferronato & Bashir 2020a, Oksanen et al. 2020, Schaefer & Straub 2016, Zhang et al. 2020).
- Extraversion: 5 papers found a positive relationship with trust in AI systems (Böckle et al. 2021, Haring et al. 2013, Kraus et al. 2020a, Merritt & Ilgen 2008,

Müller et al. 2019), while 1 paper found a negative relationship (Ferronato & Bashir 2020a). One further study found that the relationship between extraversion and trust is more complex. Elson et al. (2018) write: “Individuals high in extraversion initially rated their trust in the agent as higher than those low in extraversion. During [...] further interaction] rounds, the agent gave a recommendation contrary to what was expected in highly confident conditions [...] trust in the agent increased compared to the previous round for individual’s low in extraversion while trust in the agent decreased compared to the previous round for individuals high in extraversion [...] extraverts appeared to have their trust most dramatically affected by the conditions where the agent gave an incorrect response in the highly confident condition” (p. 436). What follows is that the level of trust in an AI system is influenced by an interaction of the degree of extraversion and whether an expectation regarding decision outcome is fulfilled.

- Conscientiousness: 3 papers found a positive relationship with trust in AI systems (Bawack et al. 2021, Chien et al. 2016, Rossi et al. 2018), while 2 papers found a negative relationship (Aliasghari et al. 2021, Oksanen et al. 2020).
- Neuroticism: 3 papers found a negative relationship with trust in AI systems (Kraus et al. 2020a, Sharan & Romano 2020, Zhang et al. 2020).

With respect to *specific personality traits*, we observed a *positive relationship* of the following constructs with trust in AI systems:

- trust propensity (Aliasghari et al. 2021, Huang & Bashir 2017, Rossi et al. 2018),
- perfect automation expectation (Lyons & Guznov 2019, Lyons et al. 2020, Matthews et al., 2020),
- dispositional trust in automation/robots (Kraus et al. 2020a, Merritt & Ilgen 2008),
- self-efficacy (Handrich 2021, Oksanen et al. 2020),
- technological innovativeness (Handrich 2021),
- sensation seeking (Zhang et al. 2020), and
- self-esteem (Kraus et al. 2020a).

Moreover, we observed a *negative relationship* between trust in AI systems and the following traits:

- internal locus of control (Chiou et al. 2021, Sharan & Romano 2020),
- negative attitude towards computers/robots (Matthews et al., 2020, Miller et al. 2021), and
- propensity to take risks (Ferronato & Bashir 2020b).

To sum up, our review of the relationship between universal personality traits and trust in AI systems indicates

that for three factors out of the Big Five (agreeableness, openness, extraversion) high values positively affect trust, while high neuroticism values negatively affect trust. With respect to conscientiousness, evidence is mixed and hence no clear statement on the relationship between this personality trait and trust in AI systems can be made based on the current research status. However, it is critical to consider that the reported results are only valid with the meaning of “all-other-things-being-equal”. As an example, an individual with high agreeableness trusts an AI system more than an individual with low agreeableness, but only if the two individuals do not differ in the other four personality traits (we reflect on this finding in the following Discussion section). Regarding the relationship between specific personality traits and trust in AI systems we found that research more frequently established a positive relationship (seven factors) than a negative relationship (three factors). However, in contrast to Big Five research the total number of available studies on most specific personality traits is lower and hence the reported results should be replicated.

Discussion and future research directions

We structure our discussion into two sub-sections. We start with a discussion of the descriptive results (as summarized in Table 1), followed by a discussion of the findings on personality traits (as summarized in Tables 2 and 3). This discussion also includes the identification of unexplored research areas and an outline of possible directions for future research.

Descriptive results

Regarding *publication year*, we identified the first studies in 2008 (Cramer et al. 2008; Merritt & Ilgen 2008). Since then, only a limited number of studies were published up until the recent past. Notably, while 29 out of the 58 identified papers were published in the period 2008-2019, the remaining 29 studies were published in 2020 and 2021. Thus, in the very recent past we observe a significant increase in empirical research at the nexus of user personality and trust in the context of AI systems, signifying the sharply rising academic interest in the topic.

Regarding *publication outlet*, we found almost a balance between conference papers (30 papers) and journal publications (27 papers). Considering that in several of the disciplines in which the investigated studies were published, conference publications are of high scientific value (e.g., computer science, HCI), we consider this balance as a “good sign”, predominantly because articles in conference proceedings guarantee that research results are quickly available (while review processes for

journals typically can take much longer). Yet, this statement should not be interpreted as an argument against journal articles. Rather, the found balance is the state which is also desirable in the future.

Regarding *scientific discipline*, we identified computer science, informatics, and robotics as the dominant group (18 papers), followed by ergonomics and human factors (13 papers), and HCI (11 papers). Information systems (IS, 8 papers) and psychology (4 papers) contributed less frequently to the current body of research. A major explanation for this observation is that the dominating disciplines have already concentrated research efforts on personality, trust, and intelligent systems in the past decade, while IS and psychology did not. However, considering that recent AI articles published in mainstream IS outlets (e.g., Berente et al., 2021, *MIS Quarterly*) explicitly refer to trust as a “managerial issue” (p. 1440), and that three of the four identified psychology papers recently appeared in a widely visible publication outlet, namely in *Frontiers in Psychology* (Kraus et al. 2020b, Miller et al. 2021, Oksanen et al. 2020), we foresee that the disciplines of IS and psychology will also contribute more research in the future. Considering that it is unlikely that the currently dominating disciplines will slow-down in their publication output (Matthews et al., 2021), it is definitive that research at the nexus of user personality and trust in the context of AI systems will most likely experience an upward tendency in the future.

Regarding *investigated systems*, our analyses revealed that robot research (including robot simulations) are by far dominant (20 papers), followed by autonomous vehicle studies (12 papers) and user interface examinations (e.g., chatbots) (11 papers). Further domains identified by our review are automated agents in the security context (e.g., airports) (7 papers), AI-driven healthcare systems (3 papers), as well as autonomous systems and AI in general (3 papers), along with some other systems such as voice assistants (3 papers). At least to some extent, this finding reflects the historical fact that computer science, informatics, and robotics, as well as ergonomics and human factors, heavily focused their research on robots (e.g., Rossi et al. 2020) and autonomous vehicles (e.g., Tenhundfeld et al. 2020). However, as a consequence of the foreseen rise of IS and psychological research, AI systems embedded into user interfaces (e.g., automated decision-support in the business context), as well as intelligent health systems and home systems (e.g., speech assistants), will likely gain in importance in the future (see, for example, a recent *MIS Quarterly* special issue, No. 3, September 2021). We explicitly make a call for studies in these under-researched domains.

Furthermore, regarding investigated systems we observed that the 58 papers remained vague in the reporting of the studies’ task instructions. It follows that an inherent

limitation of the analyzed articles is that it is *not* clear if the identified influences of personality traits on trust in AI systems are indeed specific to AI systems or hold for IT systems in general. Considering the idiosyncrasies of AI systems and their possible influences on trust (e.g., Siau & Wang, 2018; Thiebies et al., 2021), future studies should report their task instructions in detail (ideally in verbatim) in order to be better able to disentangle the specific AI effects from the more general IT effects. Such a description of task instructions should also include the definition of AI, which was provided to the study participants, if provided at all.¹⁹

Regarding applied *research method*, we observe that except for some qualitative interviews which are reported as a “by-product” of larger quantitative studies (e.g., Sarkar et al., 2017),²⁰ the current body of research almost entirely used quantitative methods. The dominant method is the laboratory experiment (22 papers), followed by survey studies (18 papers), online experiment (14 papers), and a few laboratory studies without experimental manipulation (6 papers). We see three major implications.

First, despite the uncontested value of quantitative research which is rooted in the positivist paradigm (e.g., Chen & Hirschheim, 2004), other epistemological positions and methods, particularly interpretivism and qualitative methods (e.g., Walsham, 1995), should become more important in the future. Such an interpretivist stance is often deeply rooted in a hermeneutic tradition, thereby being of a fundamentally idiographic nature. Such research (e.g., interpreting interview data on individuals’ perceptions of their interaction with AI systems and theory building), therefore, has the objective of determining “patterns” and “richness in reality” (Mason et al., 1997, p. 308).

Second, detailed analysis of the identified online experiments revealed that these studies typically used Amazon Mechanical Turk (MTurk) for data collection. MTurk is a platform that offers access to a geographically dispersed set of respondents. While such a sample “can be more representative and diverse than locally collected samples [...] respondents tend to be relatively young, digitally savvy adults” (Antes et al. 2021, p. 3). It follows that future research should replicate existing research findings with more representative populations. This call for future studies

¹⁹ The author would like to thank one anonymous reviewer for outlining this important point. To substantiate the author’s manual check of this finding, a formal search was made in all 58 papers based on the terms “task instruction” and “task description”. Two hits were found for “task instruction” (Harriott et al., 2018, Sharan & Romano 2020) and one hit was found for “task description” (Harriott et al., 2018). However, these three text passages also do not specify the task in a way so that the reader exactly knows the participants’ notion of the investigated system.

²⁰ Sarkar et al. (2017) conducted a lab experiment and as a complement collected interview data.

is substantiated by the fact that attitude towards AI systems (e.g., Nam, 2019), trust (e.g., Sutter & Kocher, 2007), and personality (e.g., Ashton & Lee, 2016) are related to age. As indicated in Appendix 2 (see table “mean age (years)”), we identified information on participant age in 60 out of the 64 studies. The average age of participants, or the dominating age group in studies in which the average age was not reported, is >40 years in only 5 studies (Matsui 2021: 44.6 y, Rossi et al. 2020: 61 y, Sorrentiono et al. 2021: 83.33 y, Voinescu et al. 2018: 67.52 y, Youn & Jin 2021: 40.8 y). Therefore, the findings of the current review are predominantly valid for younger age groups. Future research should be based more often on older people. This call is substantiated by the fact that older people are increasingly concerned by AI systems, as signified by the example of socially assistive robots (e.g., Sorrentiono et al. 2021).

Third, during the paper screening and selection process (see Fig. 2) we encountered a paper by Mühl et al. (2020) who used neurophysiological measurement (i.e., skin conductance) as a complement to self-reports. However, because this paper studied passengers’ preferences and trust while being driven by a human driver and an intelligent vehicle without any focus on user personality, we did *not* consider this article in our body of analyzed literature. It follows that no single article in our sample of N = 58 papers used neurophysiological measurement. Importantly, research has demonstrated “associations between personality traits and the structure and function of the nervous system” (Funder, 2001, p. 206). Thus, we consider the application of neuroscientific measurement techniques as a promising methodological avenue for future studies. In this respect, the study by Mühl et al. (2020) could be a promising starting point. This call for the use of neuroscience approaches is substantiated by two research developments: first, the increasingly important role of neuroscience approaches in IS research, referred to as NeuroIS (e.g., Dimoka et al., 2012; Riedl & Léger, 2016), and second, the increasing evidence that personality traits have a strong evolutionary, affective, and hence biological component (e.g., Montag et al., 2016; Montag & Panksepp, 2017).

Regarding *country of data collection*, the most striking result is that in 30 out of 60 studies in which the country is reported, data gathering took place in the United States. Then, the category “mix of different countries” follows. This category comprises MTurk studies in which subjects from more than one country participated. Six studies collected data in Germany, followed by the UK (4 studies), Canada (3 studies), Italy, Australia, and Japan (2 studies each), as well as three further countries in which one study was conducted (China, Korea, South Africa). The major implication of this finding is that more future studies should be carried out outside the United States. This call is substantiated by evidence showing that personality and culture are related (e.g., Leung

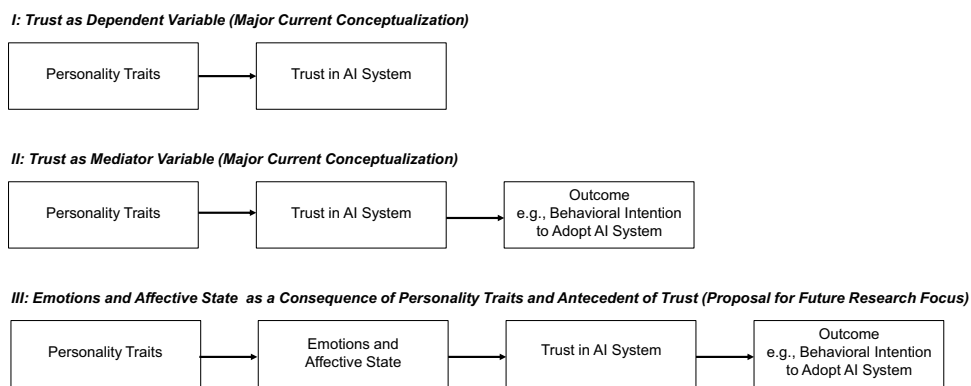
& Cohen, 2011; McCrae, 2000). One study in our sample investigated the relationship between trust attitudes towards automated digital technologies, Hofstede’s cultural dimensions, and the Big Five personality traits (Chien et al. 2016). Reflecting on their results, Chien et al. write: “the U.S. population had the highest trust score and Turkish group scored the lowest, with Taiwanese population falling in between [...] Evaluations of the inter-relational aspects of personality and general trust showed that an individual with a high trait of agreeableness or conscientiousness had increased trust in automation” (p. 845). Thus, culture should be a construct of interest in future research. This call for future studies is substantiated by long existing calls for more cultural studies in a seminal psychology paper. Funder (2001) writes: “[the] direction is to try to distinguish between the psychological elements that are shared by all cultures (etics) and those that are distinctive to particular cultures (emics) [...] The big five have been offered as possible etics” (p. 203). Against this background, fruitful avenues for future research could be to examine the possible relationship between culture and personality traits in the context of trust in AI systems.

A seminal paper by Hofstede and McCrae (2004) could serve as a starting point.²¹ They found that several culture dimensions are related to the Big Five; specifically, they report the following statistically significant correlations²²: *power distance* (“the extent to which the less powerful members of organizations and institutions [...] accept and expect that power is distributed unequally”) correlates with extraversion (–), conscientiousness (+), and openness (–); *uncertainty avoidance* (“a society’s tolerance for ambiguity [...] indicating] to what extent a culture programs its members to feel either uncomfortable or comfortable in unstructured situations [...] which] are novel, unknown, surprising, and different than usual”) correlates with neuroticism (+) and agreeableness (–); *individualism* (“the degree to which individuals are integrated into groups [...] and in] individualist societies, the ties between individuals are loose: everyone is expected to look after himself or herself and his or her immediate family”) correlates with extraversion (+); finally, *masculinity* (“the distribution of emotional roles between the sexes [...] women’s values differ less among societies than men’s values [...] and] men’s values vary along a dimension from very assertive and competitive and maximally different from women’s values on one side to modest and caring and similar to women’s values on the other”) correlates with

²¹ The platform www.hofstede-insights.com defines culture as “the collective mental programming of the human mind which distinguishes one group of people from another” and it offers a feature to compare countries based on various culture dimensions.

²² Definitions of culture dimensions are taken in verbatim from Hofstede and McCrae (2004, pp. 62-63). The signs “+” and “–” indicate positive or negative correlations.

Fig. 3 Two dominating theoretical conceptualizations of personality and trust in the context of AI system adoption and proposal for future research focus



openness (+), neuroticism (+), and agreeableness (–). This existing knowledge should be used in future studies on trust in AI systems.

Findings on personality traits

Regarding the *universal personality traits*, we found that the Big Five were studied extensively, while other general personality models were not. Thus, future research should consider these other models more frequently. In particular, we recommend future studies based on HEXACO (Ashton et al., 2004) and NEO-PI-3 (McCrae et al., 2005), as these models are both related to the Big Five and relatively novel frameworks in the personality literature (if compared to Myers-Briggs’ and Eysenck’s models). In our sample of analyzed papers, Müller et al. (2019) used HEXACO to study personality and trust in the context of human-chatbot interaction. Moreover, Rossi et al. (2020) used NEO-PI-3 to examine human interaction with socially assistive robots in a simulated home environment. These two studies may serve as a starting point for future research.

Regarding the *theoretical conceptualization* of personality and trust in studies on adoption and interaction with AI systems (see Table 3), our review reveals two dominating causal chains. First, personality is conceptualized as an independent variable and trust in the AI system as a dependent variable (e.g., Elson et al. 2020). Second, personality is conceptualized as an independent variable, trust in the AI system as a mediator variable, and another outcome such as behavioral intention to adopt the AI system is conceptualized as a dependent variable (Handrich 2021). Figure 3 graphically summarizes these two theoretical mechanisms, and we added a third one. This additional mechanism describes a fruitful future research focus in which emotions and affective state is conceptualized as a consequence of personality traits and as an antecedent of trust in AI system, which in turn influences behavioral intention. A major motivation for consideration of this additional mechanism is the fact that emotions and affect have become critical constructs in IS

research in general (e.g., Zhang, 2013) and in AI system adoption research in particular (e.g., Kraus et al. 2020b, Miller et al. 2021).

The Merriam-Webster Dictionary defines *affect* as “a set of observable manifestations of an experienced emotion: the facial expressions, gestures, postures, vocal intonations, etc., that typically accompany an emotion”.²³ Thus, emotions are closely related to physiological processes, followed by observable consequences of these processes—the affect. Interestingly, despite a few notable exceptions such as Kraus et al.’s (2020b) work on automated driving or Miller et al.’s (2021) work on human-robot interaction, emotions and affect hardly play a role in the analyzed literature. Thus, the role of emotions and affect has hardly been established in the present research context.

In Fig. 3 (bottom), we suggest positioning emotions and affective state as a consequence of personality traits and as an antecedent of trust, which in turn influences further downstream variables, such as behavioral intention to adopt an AI system. This positioning is consistent with the theoretical frameworks in Kraus et al. (2020b) and Miller et al. (2021). For example, Miller et al. (2021) write: “Besides user dispositions, users’ emotional states during the familiarization with a robot are a potential source of variance for robot trust. As the experience of emotional states has been shown to be considerably affected by personal dispositions, this research proposes a general mediation mechanism from the effects of user dispositions on trust in automation by user states [...] focus[ing] on state anxiety as a specific affective state, which is expected to explain interindividual differences in trust in robots [...] State anxiety is defined as “subjective, consciously perceived feelings of apprehension and tension, accompanied by or associated with activation or arousal of the autonomic nervous system” [...]” (p. 5). Future empirical research should directly test this causal

²³ <https://www.merriam-webster.com/dictionary/affect> (accessed on January 16, 2022).

chain.²⁴ Methodologically, we recommend two things: first, to complement existing personality measurement instruments (e.g., Big Five) with the Affective Neuroscience Personality Scales (ANPS) (Montag et al., 2021; Orri et al., 2017; Reuter et al., 2017), and second, to consider measuring emotions based on neurophysiological measurement (NeuroIS). The major argument for this suggestion is that trusting beliefs in and attitudes towards AI systems are not only influenced by conscious perceptions and thoughts (which would imply the use of self-report measurement). Rather, they are also influenced by unconscious perceptions and information processing which cannot be captured based on self-reports alone (e.g., Dimoka et al., 2012; Riedl & Léger, 2016; vom Brocke et al., 2020). Thus, future research should consider measuring emotions, and measurement should include both self-report and neuroscience instruments.

As outlined in detail in the Results section, we examined the *Big Five's influence on trust in AI systems*. In essence, overwhelming evidence shows that both agreeableness and openness positively affect trust. Moreover, while the evidence is less compelling if compared to agreeableness and openness, there is still a clear tendency that extraversion also positively affects trust. Regarding neuroticism, the findings are also clear-cut. Less neurotic people exhibit more trust in AI systems. However, this finding is based on fewer studies if compared to the other Big Five traits.

Finally, regarding the influence of conscientiousness on trust in AI systems the evidence is mixed. While some studies (Bawack et al. 2021, Chien et al. 2016, Rossi et al. 2018) found a positive relationship, other studies found a negative relationship (Aliasghari et al. 2021, Oksanen et al. 2020). For example, based on the context of a domestic robot for food preparation, Aliasghari et al. (2021) found that more conscientious people trusted the robot less, and hence were more likely to rely on themselves as cook or a restaurant. In contrast, another study which was carried out in the context of AI-based voice shopping (e.g., Amazon Echo) found a positive relationship between conscientiousness and trust (Bawack et al. 2021). What follows is that variance in AI system (domestic robot vs. AI-based voice assistant) and/or in the task (food preparation vs. shopping), may have caused the different relationships between the personality trait conscientiousness and trust. Thus, it is a fruitful scientific endeavor to study more systematically the influence of specific AI systems and/or the technology-supported tasks (as moderators) on the relationship between personality (independent variable) and trust (dependent variable).

We highlight that trust can vary strongly depending on what kind of AI system a user is dealing with. For example, highly critical AI systems such as medical AI systems would be harder to trust than non-critical systems such as voice assistants or recommender systems (due to the potential harm they might cause in case of a malfunction or a poor system design). Thus, despite the fact that the current research status suggests a direct relationship between the personality traits agreeableness (positive), openness (positive), extraversion (positive), and neuroticism (negative) and trust in AI systems, more research is necessary to determine whether these findings hold in the context of highly critical AI systems. Importantly, no paper in our sample of 58 empirical articles chose the context of highly critical AI systems like those in the medical domain. Thus, today we do not know whether the current results may be generalized to this context. Also, other boundary conditions should be considered in future research, including user experience and computer self-efficacy (Matthews et al., 2021).

Moreover, the kind of collaboration process should be considered in future studies. In essence, decision automation should be distinguished from scenarios where AI acts as a decision support system. Our analysis of the extant literature (see Appendix 2, column “Context”) indicates that the available literature predominantly dealt with decision support systems rather than complete decision automation. Thus, the influence of personality traits on trust in AI systems which fully automate decisions should be studied more frequently in the future. Several researchers are aware of this necessary distinction; however, they have not yet considered this distinction in their research designs. As an example, Tenhundfeld et al. (2020) studied trust in automated parking based on a Tesla vehicle. However, they highlight that their research focus was “*partially* automated parking” (p. 194, italics added) and not complete decision automation.²⁵

Despite the fact that current knowledge on the relationship between the Big Five traits and trust in an AI system is substantial, the practical implication of this knowledge is limited. Every human can be characterized as a configuration of the Big Five values. In the simplest case in which each trait is assumed to be either high or low we have $2^5 = 32$ personality types. However, the existing body of literature does *not* focus on investigation of these types' influence on trust. Rather, the extant literature established a trait's individual relationship with trust in a specific AI system. This kind of knowledge is limited as it only allows for trust prediction for some configurations. Based on the present review's knowledge, a person with profile₁ would

²⁴ Kraus et al. (2020b) theoretically suggest *and* empirically confirm the following causal chain: personality traits as independent variable, emotional state as mediator, and trust as dependent variable. Replication studies should be conducted to better establish this mechanism.

²⁵ The author would like to thank one anonymous reviewer for outlining this important point.

trust an AI system more than a person with profile₂ (initial letters of the Big Five, “↑” indicates a high level, while “↓” indicates a low level): profile₁ = A↑ + O↑ + E↑ + N↓ + C↑, profile₂ = A↓ + O↓ + E↓ + N↑ + C↑. However, what would be the prediction for the following configuration: profile₃ = A↑ + O↓ + E↑ + N↑ + C↑? Based on the insights presented in the extant literature, it would be hardly possible to make a sound prediction. However, one paper in the analyzed literature used latent profile analysis to study trust in chatbots (e.g., Alexa). Based on this technique and the HEXACO model, Müller et al. (2019) not only identified extraversion, agreeableness, and honesty-humility as the three most relevant traits for trust prediction, but also determined *specific personality profiles with predictive power for trust*: “introverted, careless, distrusting user”, “conscientious, curious, trusting user”, and “careless, dishonest, trusting user” (p. 35). More research of this sort is necessary in the future.

Regarding *specific personality traits*, we found some clear-cut results. First, trust propensity (Aliasghari et al. 2021, Huang & Bashir 2017, Rossi et al. 2018), perfect automation expectation (Lyons & Guznov 2019, Lyons et al. 2020, Matthews et al., 2020), and dispositional trust in automation/robots (Kraus et al. 2020a, Merritt & Ilgen 2008) positively affect situational trust in a specific AI system. Moreover, we found the following traits to also positively affect this trust: self-efficacy, technological innovativeness, sensation seeking, and self-esteem. However, as indicated in the Results section, because only a very limited number of studies have examined these constructs, future research should replicate these findings before more definitive conclusions can be drawn.

Two specific traits turned out to be negative predictors of trust in AI systems (at least two independent studies observed this effect for both factors): negative attitude towards computers/robots (Matthews et al., 2020, Miller et al. 2021) and internal locus of control (Chiou et al. 2021, Sharan & Romano 2020). It is highly plausible that people with a negative attitude towards computers and robots generally do exhibit little situational trust in a specific AI system. Regarding locus of control, evidence shows that people who believe that they can control their own lives (rather than being controlled by external forces) exhibit less trust towards AI systems. This finding is also plausible because a strong belief in one’s own ability to control life may come along with less trust in everything else which could control one’s life, either other humans or technological artifacts like AI-based systems. The study by Choi & Lee (2015) provides another interesting rationale for this finding; based on an autonomous vehicle context they write: “locus of control [...] selected as driving-related personality factor [...] External locus of control significantly influenced behavior. This result demonstrated that someone who experiences difficulty

in driving, such as older adults, has greater intention to use the autonomous vehicles” (p. 699). It follows that people who believe that they do not have sufficient capabilities and skills to interact with or handle an intelligent system (= no internal locus of control) do trust an intelligent system *because* there is no alternative.

Overall, a major conclusion of the current review is that today only a limited number of specific personality traits were studied sufficiently so that solid knowledge exists (i.e., trust propensity, perfect automation expectation, and dispositional trust in automation/robots positively affect trust in AI systems, while negative attitude towards computers/robots and internal locus of control have a negative effect). However, more research should be conducted on the remaining traits which we found in our literature analysis (as described in detail in the column “Specific traits” in Appendix 2) as well as further traits which are described elsewhere (Matthews et al., 2021).

Our call for more research on specific personality traits is substantiated by the fact that the Big Five do *not* subsume “all there is to say about personality” (Funder, 2001, p. 200). As an example, consider an authoritarian personality (Adorno et al., 1950). Funder (2001) indicates that such a personality “would be high on conscientiousness and low on agreeableness and openness [...] but much would be lost if we tried to reduce our understanding of authoritarianism to these three dimensions” (p. 201). It follows that while specific personality traits are definitely related to more abstract traits such as the Big Five, they feature additional information. Therefore, we reiterate for the domain of trust in AI systems what Funder (2001) already declared more than two decades ago with reference to the Big Five, namely that researchers should not be “seduced by convenience [...] and] act as if they can obtain a complete portrait of personality by grabbing five quick ratings” (p. 201).

As indicated in Table 3 and Fig. 3 (see “Trust as Mediator”), some studies used behavioral intention, or adoption intention, as dependent variable. Thus, these studies provide first insights on how personality analysis can be used to improve the adoption process of AI systems. One study, for example, indicates that system transparency (e.g., “I believe that I can form a mental model and predict future behavior of the autonomous vehicle.”), technical competence (e.g., “I believe that I can depend and rely on the autonomous vehicle.”), and situation management (e.g., “I believe that the autonomous vehicle will provide adequate, effective, and responsive help.”) positively affected trust, which, in turn, increased behavioral intention to use an autonomous vehicle (Choi & Ji, 2015). In another study, it is reported that need for interaction, self-efficacy, technological innovativeness, and novelty seeking positively affected trust, which, in turn, reduced innovation resistance and increased adoption intention of intelligent personal assistants like Alexa (Handrich

2021). These and further studies (Cohen & Sergay 2011, Hegner et al. 2019, Zhang et al. 2020) constitute a starting point for future research that has the goal to develop recommendations on how to improve the adoption process of AI systems based on user personality knowledge. As signified by the examples above, current research suggests—if considered as collective evidence—that the *interaction* of personality traits (either the Big Five or specific traits like those investigated by Handrich 2021) with system properties (like those investigated by Choi & Ji, 2015) determines trust. In addition to system properties, task demands (e.g., workload in partly automated tasks) should also be considered. This need is substantiated by Szalma and Taylor’s (2011) work—based on significant empirical evidence, they write: “[T]he impact of automation and task demand on participant response depends to a substantial degree on human personality traits. A practical implication is that the identification of specific profiles to predict human–automation interaction is unlikely to be useful as a generic selection measure [...] the complexity of the interactive effects suggest[s] that for selection tools to be effective, the optimal trait profiles would need to be developed separately for each level, type, and perhaps even domain of application of automation” (p. 91). We make a call for future studies on such interaction effects and their consequences for adoption decisions.

Design implications and adaptive systems as possible future design science research focus

Design science research has evolved as a major academic paradigm in the IS discipline, which aims to design innovative and useful IT artifacts (e.g., Hevner et al., 2004; March & Smith, 1995; Peffers et al., 2008). A major question in the current research context is how the presented empirical research can inform the design of AI systems in order to increase trust. In essence, though the 58 reviewed works have made valuable contributions to the literature (predominantly from a theoretical and empirical point of view, i.e., descriptive knowledge), the authors did not have an explicit intent to systematically conduct their research as design science research projects, nor was that the result. What follows is that *concrete* design implications (i.e., a form of prescriptive knowledge) of the current body of knowledge are limited.²⁶ This finding of the present review is consistent with a recent observation by Ahmad et al. (2022) who indicate that “previous studies mainly employ empirical methods to describe behaviors in interaction with these systems and consequently generate descriptive knowledge about the use and effectiveness of personality-adaptivity CAs [Conversational Agents,

²⁶ This paucity of design focus is substantiated by the fact that only one out of the 58 analyzed papers uses the term “design” in its title (Voinescu et al. 2018).

a specific type of AI system]. In addition [...] there is still a research gap in providing prescriptive knowledge about how to design a PACA [personality-adaptive conversational agent] to improve interactions with users” (p. 3).

However, the analyzed literature offers some *abstract* design implications. Yet, these implications typically do not go beyond the recommendations which are already documented in existing conceptual papers or review works (e.g., Siau & Wang, 2018). Obviously, this constitutes a major limitation of the current body of knowledge which should be addressed in future research. As an example, in one of our 58 analyzed papers, Aliasghari et al. (2021) found, among other things, that “the participants lost trust in the robot [...] when the robot made an error after it seemed to have learned the task by making no mistakes [...] and] personality traits of the participants [...] affected some aspects of their trust in the robot” (p. 87). A design recommendation that would result from such an empirical finding is trivial: Developers should design AI systems which do not make errors. In a conceptual paper, Siau and Wang (2018) already indicated several years before the publication of the Aliasghari et al.’s study that trust “must be nurtured and maintained [...] which] happens through [...] competence of AI in completing tasks and finishing tasks in a consistent and reliable manner [...] and] there should be no unexpected downtime or crashes” (p. 51).²⁷ Against the background of our finding that concrete design recommendations are hardly available in the current body of knowledge, we make a call for future IS design science research in the context of user personality and trust in AI systems. Then, more concrete design recommendations can be expected in the future.

As a starting point for future research, consider the following example within the mentioned domain of system errors. Foremost, it is *not* possible to create AI systems that do not make errors, independent from whether they are classical knowledge-based systems, are trained using machine learning algorithms, or are hybrids.²⁸ The aim must therefore be to prepare the user that errors might occur and how to cope with it (e.g., by implementing cross-check procedures or using an explanation component) (e.g., Shin, 2021). Also, expectation management is critical. Hence, it is important that the claims coming with an AI system are realistic because the gap between expectations and perceived delivery predominantly

²⁷ Note that Siau & Wang (2018, pp. 50-52) describe several other factors which designers could consider in order to foster trust. Examples are representation (e.g., “a robot dog is another example of an AI representation that humans find easier to trust. Dogs are human’s best friends and represent loyalty and diligence”) or transparency and explainability (e.g., “challenges in machine learning and deep learning is the black box in the ML and decision-making processes. If the explainability of the AI application is poor or missing, trust is affected”), among several other factors (e.g., usability and reliability, collaboration and communication, sociability and bonding, security and privacy protection).

²⁸ The author would like to thank one anonymous reviewer for outlining this important point.

determines perceived quality and related trust perceptions (e.g., Jiang et al., 2002). However, today neither does scientific knowledge exist on how specific personality types perceive different preparation procedures, coping strategies, or expectation management approaches, nor do we know the procedures', strategies', and approaches' efficacy regarding trust as a function of user personality. This research gap should be closed in the future. Szalma and Taylor's (2011, p. 93) "general guidelines for incorporating individual differences into the design of human-technology interfaces" could serve as a starting point for corresponding design science research.

Moreover, a future design science research program should explicitly consider the development of adaptive systems. Recently, Matthews et al. (2021) substantiated the need for adaptive systems. They argued that task analysis is "a crucial step in the design of any system and its interfaces [... and] there should be an analogous 'person analysis'"—moreover, they foresee "individuation in design" and "adaptive technologies" because "conventional 'static' products offer little or no flexibility to accommodate variation in personality and individual differences once designed"

(p. 7). Against this background, future studies should contribute to the conceptualization and development of such adaptive systems. Based on publications in the literature on real-time adaptive interventions (for an overview paper, see Nahum-Shani et al., 2018), we define such adaptive systems as a technology aiming to support the execution of a task by adapting to a user's personality, as well as to resulting neurophysiological states and behavior.

Imagine a situation in which a user interacts with an AI system. The AI system could be a chatbot or recommender system which is presented on the computer screen. In principle, however, the basic idea of adaptive systems also applies to other AI systems such as robots or autonomous vehicles. First, the system must be capable of recording the user state (e.g., emotions), its own state, context, and behavioral user data (e.g., preferences via system input in recommender systems). Second, the system must integrate this stream of data to derive the user's personality profile (either universal personality traits like the Big Five or specific personality traits like propensity to take risks or computer anxiety). Third, based on the personality profile the AI system adapts in real-time in

Table 4 Summary of example works in which AI systems adapt based on user personality

Study	Focus and major result
Ahmad et al. (2022)	The authors argue that currently conversational agents (CAs), a specific type of AI system, do not live up to their full potential because they are unable to provide responses tailored to users' personalities. To address this problem, the authors conducted a design science research project with the goal to design personality-adaptive conversational agents (PACAs). Based on an iterative and multi-step approach, they derive and formulate six design principles for PACAs for the domain of mental health care. The findings of the evaluation with psychologists and psychiatrists suggest that PACAs can be a promising source of mental health support. The authors indicate that with their design principles, they contribute to the body of design knowledge for CAs and provide guidance for practitioners who intend to design PACAs.
Kampman et al. (2019)	The authors describe a virtual agent whose major function is adaptation to user personality. User personality is determined through two approaches: audio data and speech transcription text. The virtual agent was implemented and evaluated. Based on a Convolutional Neural Network, both approaches show good performance. The authors conclude that perception of user personality is possible using audio and text data. Moreover, the authors report correlations between user personality and preferred agent personality. Specifically, results suggest that the user personality openness correlates with a preference for agents with more gentle personality. Also, it was found that users report more empathy and enjoyed conversations more when agents adapted to their personality.
Tapus et al. (2008)	The authors describe a fully automated socially assistive therapist robot. The robot is able to monitor, assist, encourage, and socially interact with post-stroke patients in a rehabilitation context. The authors show that it is possible to develop a behavior adaptation system. Specifically, based on validation experiments it was demonstrated that the robot is able to adjust its social interaction parameters such as interaction distances, speed, or vocal content based on the user's personality and task performance. Moreover, it was demonstrated that adaptation to the user's personality may lead to improved human task performance.
Yorita et al. (2019)	Based on empirical evidence, the authors argue that people relate better with other people who have similar personality characteristics and that this fact is useful in scenarios where people should be receptive to receiving support and advice from others. The authors describe a chatbot personality model and a genetic algorithm that enables the chatbot to adapt its personality in real-time as it interacts with the user. The model draws upon the Big Five framework and focuses on extroversion and agreeableness. Results of the evaluation study indicate that the algorithms are capable of adapting personality traits to match the identified traits of the user.
Zalake (2020)	Interventions based on human interaction with virtual agents have been used to change user health-related attitudes and behaviors. The authors argue that current virtual agents use several persuasion techniques to influence user behavior. However, available technologies typically do not account for individual differences between users. This is problematic, because individual differences between users significantly affect the efficacy of different persuasion strategies. The study proposes a persuasive virtual agent-based mobile health intervention that adapts to individual differences in user personality to promote anxiety coping strategies.

order to maximize the user's trust in the system (foundations of such adaptive systems can be found, for example, in Byrne & Parasuraman, 1996, Harriott et al., 2018, or Picard, 1997).

It is critical to stress that research has already demonstrated the feasibility of adaptive systems based on user personality. Table 4 provides a summary of example works which could serve as a starting point for future design science research projects. Thus, despite the challenges which come along with the development of adaptive systems (e.g., Adam et al., 2017; Picard, 2003; vom Brocke et al., 2020), we foresee a prosperous development of research in this important domain. This positive outlook is substantiated by recent research which has demonstrated the feasibility of unobtrusive real-time measurement of trust in machines based on physiological measurement (Akash et al., 2018; Walker et al., 2019). Importantly, the development of such adaptive systems must not ignore the legal, ethical, organizational, and social aspects, which need to be explored in future studies. Because the development of adaptive systems is a highly interdisciplinary phenomenon, we expect a number of different scientific disciplines to contribute to this research (e.g., IS, informatics, organization science, and psychology). However, considering the IS discipline's successful history in interdisciplinary research endeavors, IS researchers could take a lead role in this area.

Limitations and concluding statement

In the present article, we reviewed 58 papers published in the period 2008–2021. To the best of our knowledge, this analysis is the first systematic review of the scientific literature on the relationship between personality traits and trust in the context of AI systems. Yet, the present review has limitations: First, while we can safely assume that the literature basis of the current review is extensive (because we searched for literature with several keywords and in four databases, including several hundred relevant journals and conference proceedings), there is no formal measure to prove completeness. Second, while we consider our literature collection and analysis approaches as rigorous, our interpretation of the findings, and especially the formulation of future research implications, is predominantly interpretive. Yet, because we report the analyzed literature in a very detailed way (see also Appendix 2), other researchers can directly draw upon our basis of 58 papers to complement or revise the current interpretations. What follows is that while we consider our review as a systematic documentation of the

research status on personality and trust in the context of AI systems, it is hoped that this article instigates further research, thereby being more of a sort of “starting point” rather than the “final word”. It will be rewarding to see what insights future research will reveal.

Appendix 1

The following list summarizes the 58 references which constitute the basis of the literature review.

- 1 Aliasghari, P., Ghafurian, M., Nehaniv, C. L., & Dautenhahn, K. (2021). Effect of domestic trainee robots' errors on human teachers' trust. *2021 30th IEEE International Conference on Robot and Human Interactive Communication, RO-MAN 2021*, 81–88. <https://doi.org/10.1109/RO-MAN50785.2021.9515510>.
- 2 Antes, A. L., Burrous, S., Sisk, B. A., Schuelke, M. J., Keune, J. D., & DuBois, J. M. (2021). Exploring perceptions of healthcare technologies enabled by artificial intelligence: an online, scenario-based survey. *BMC Medical Informatics and Decision Making*, 21(1), 221. <https://doi.org/10.1186/s12911-021-01586-8>.
- 3 Bawack, R. E., Wamba, S. F., & Carillo, K. D. A. (2021). Exploring the role of personality, trust, and privacy in customer experience performance during voice shopping: Evidence from SEM and fuzzy set qualitative comparative analysis. *International Journal of Information Management*, 58, 102309. <https://doi.org/10.1016/j.ijinfomgt.2021.102309>.
- 4 Böckle, M., Yeboah-Antwi, K., & Kouris, I. (2021). Can you trust the black box? The effect of personality traits on trust in AI-enabled user interfaces. In D. H. & N. S. (Eds.), *Artificial intelligence in HCI. HCII 2021. Lecture Notes in Computer Science* (Vol. 12797, pp. 3–20). Springer. https://doi.org/10.1007/978-3-030-77772-2_1.
- 5 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>.
- 6 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>.

- 7 Choi, J. K., & Ji, Y. G. (2015). Investigating the importance of trust on adopting an autonomous vehicle. *International Journal of Human-Computer Interaction*, 31(10), 692–702. <https://doi.org/10.1080/10447318.2015.1070549>.
- 8 Cohen, J. F., & Sergay, S. D. (2011). An empirical study of health consumer beliefs, attitude and intentions toward the use of self-service kiosks. *17th Americas Conference on Information Systems 2011, AMCIS 2011 Proceedings - All Submissions*, 46, 403–412.
- 9 Conati, C., Barral, O., Putnam, V., & Rieger, L. (2021). Toward personalized XAI: A case study in intelligent tutoring systems. *Artificial Intelligence*, 298, 103503. <https://doi.org/10.1016/j.artint.2021.103503>.
- 10 Cramer, H., Evers, V., Kemper, N., & Wielinga, B. (2008). Effects of autonomy, traffic conditions and driver personality traits on attitudes and trust towards in-vehicle agents. *2008 IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology*, 477–482. <https://doi.org/10.1109/WIIAT.2008.326>.
- 11 Elson, J. S., Derrick, D., & Ligon, G. (2018). Examining trust and reliance in collaborations between humans and automated agents. *Proceedings of the 51st Hawaii International Conference on System Sciences*, 430–439. <https://doi.org/10.24251/HICSS.2018.056>.
- 12 Elson, J. S., Derrick, D., & Ligon, G. (2020). Trusting a humanoid robot: exploring personality and trusting effects in a human-robot partnership. *Proceedings of the 53rd Hawaii International Conference on System Sciences*, 543–552. <https://doi.org/10.24251/HICSS.2020.067>.
- 13 Ferronato, P., & Bashir, M. (2020a). An examination of dispositional trust in human and autonomous system interactions. In K. M. (Ed.), *human-computer interaction. Human values and quality of life. HCII 2020. Lecture Notes in Computer Science* (Vol. 12183, pp. 420–435). Springer. https://doi.org/10.1007/978-3-030-49065-2_30.
- 14 Ferronato, P., & Bashir, M. (2020b). Does the propensity to take risks influence human interactions with autonomous systems? In C. I., N. E., & A. T. (Eds.), *advances in human factors in cybersecurity. AHFE 2020. Advances in Intelligent Systems and Computing* (Vol. 1219, pp. 23–29). Springer. https://doi.org/10.1007/978-3-030-52581-1_4.
- 15 Gibson, A. M., Alarcon, G. M., Jessup, S. A., & Capiola, A. (2020). “Do you still trust me?” effects of personality on changes in trust during an experimental task with a human or robot partner. *Proceedings of the 53rd Hawaii International Conference on System Sciences*, 5099–5108.
- 16 Gillath, O., Ai, T., Branicky, M. S., Keshmiri, S., Davison, R. B., & Spaulding, R. (2021). Attachment and trust in artificial intelligence. *Computers in Human Behavior*, 115, 106607. <https://doi.org/10.1016/j.chb.2020.106607>.
- 17 Handrich, M. (2021). Alexa, you freak me out – identifying drivers of innovation resistance and adoption of intelligent personal assistants. *ICIS 2021 Proceedings*, 11, 1–17.
- 18 Hanna, N., & Richards, D. (2015). The influence of users’ personality on the perception of intelligent virtual agents’ personality and the trust within a collaborative context. In F. K. et Al. (Ed.), *CARE-MFSC 2015, CCIS* (Vol. 541, pp. 31–47). Springer. https://doi.org/10.1007/978-3-319-24804-2_3.
- 19 Haring, K. S., Matsumoto, Y., & Watanabe, K. (2013). How do people perceive and trust a lifelike robot. *Proceedings of the World Congress on Engineering and Computer Science 2013, 1*, 425–430.
- 20 Harriott, C. E., Garver, S., & Cunha, M. (2018). A motivation for co-adaptive human-robot interaction. In C. J. (Ed.), *advances in human factors in robots and unmanned systems, advances in intelligent systems and computing* (Vol. 595, pp. 148–160). Springer. https://doi.org/10.1007/978-3-319-60384-1_15.
- 21 Hegner, S. M., Beldad, A. D., & Brunswick, G. J. (2019). In automatic we trust: investigating the impact of trust, control, personality characteristics, and extrinsic and intrinsic motivations on the acceptance of autonomous vehicles. *International Journal of Human-Computer Interaction*, 35(19), 1769–1780. <https://doi.org/10.1080/10447318.2019.1572353>.
- 22 Huang, H.-Y., & Bashir, M. (2017). Personal influences on dynamic trust formation in human-agent interaction. *Proceedings of the 5th International Conference on Human Agent Interaction*, 233–243. <https://doi.org/10.1145/3125739.3125749>.
- 23 Huang, H.-Y., Twidale, M., & Bashir, M. (2020). ‘If you agree with me, do i trust you?’: an examination of human-agent trust from a psychological perspective. In B. R. Bi Y. & K. S. (Eds.), *intelligent systems and applications. IntelliSys 2019. Advances in Intelligent Systems and Computing* (Vol. 1038, pp. 994–1013). Springer. https://doi.org/10.1007/978-3-030-29513-4_73.

- 24 Kim, K. J., Park, E., Sundar, S. S., & del Pobil, A. P. (2012). The effects of immersive tendency and need to belong on human-robot interaction. *Proceedings of the Seventh Annual ACM/IEEE International Conference on Human-Robot Interaction - HRI '12*, 207–208. <https://doi.org/10.1145/2157689.2157758>.
- 25 Kim, W., Kim, N., Lyons, J. B., & Nam, C. S. (2020). Factors affecting trust in high-vulnerability human-robot interaction contexts: A structural equation modelling approach. *Applied Ergonomics*, 85, 103056. <https://doi.org/10.1016/j.apergo.2020.103056>.
- 26 Kraus, J., Scholz, D., & Baumann, M. (2020a). What's driving me? exploration and validation of a hierarchical personality model for trust in automated driving. *Human Factors*, 63(6), 1076–1105. <https://doi.org/10.1177/0018720820922653>.
- 27 Kraus, J., Scholz, D., Messner, E.-M., Messner, M., & Baumann, M. (2020b). 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>.
- 28 Lyons, J. B., & Guznov, S. Y. (2019). Individual differences in human-machine trust: A multi-study look at the perfect automation schema. *Theoretical Issues in Ergonomics Science*, 20(4), 440–458. <https://doi.org/10.1080/1463922X.2018.1491071>.
- 29 Lyons, J. B., Nam, C. S., Jessup, S. A., Vo, T. Q., & Wynne, K. T. (2020). The role of individual differences as predictors of trust in autonomous security robots. *2020 IEEE International Conference on Human-Machine Systems (ICHMS)*, 1–5. <https://doi.org/10.1109/ICHMS49158.2020.9209544>.
- 30 Matsui, T. (2021). Relationship between users' trust in robots and belief in paranormal entities. *Proceedings of the 9th International Conference on Human-Agent Interaction*, 252–256. <https://doi.org/10.1145/3472307.3484666>.
- 31 Matthews, G., Lin, J., Panganiban, A. R., & Long, M. D. (2020). Individual differences in trust in autonomous robots: implications for transparency. *IEEE Transactions on Human-Machine Systems*, 50(3), 234–244. <https://doi.org/10.1109/THMS.2019.2947592>.
- 32 McBride, M., Carter, L., & Ntuen, C. (2012). The impact of personality on nurses' bias towards automated decision aid acceptance. *International Journal of Information Systems and Change Management*, 6(2), 132–146. <https://doi.org/10.1504/IJISCM.2012.051148>.
- 33 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>.
- 34 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>.
- 35 Miller, L., Kraus, J., Babel, F., & Baumann, M. (2021). More than a feeling—interrelation of trust layers in human-robot interaction and the role of user dispositions and state anxiety. *Frontiers in Psychology*, 12, 592711. <https://doi.org/10.3389/fpsyg.2021.592711>.
- 36 Müller, L., Mattke, J., Maier, C., Weitzel, T., & Graser, H. (2019). Chatbot acceptance: A latent profile analysis on individuals' trust in conversational agents. *SIGMIS-CPR 2019 - Proceedings of the 2019 Computers and People Research Conference*, 35–42. <https://doi.org/10.1145/3322385.3322392>.
- 37 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>.
- 38 Perelman, B. S., Evans, A. W., & Schaefer, K. E. (2017). Mental model consensus and shifts during navigation system-assisted route planning. *Proceedings of the Human Factors and Ergonomics Society 2017 Annual Meeting*, 61(1), 1183–1187. <https://doi.org/10.1177/1541931213601779>.
- 39 Pop, V. L., Shrewsbury, A., & Durso, F. T. (2015). Individual differences in the calibration of trust in automation. *Human Factors*, 57(4), 545–556. <https://doi.org/10.1177/0018720814564422>.
- 40 Rossi, A., Dautenhahn, K., Koay, K. L., & Walters, M. L. (2018). The impact of peoples' personal dispositions and personalities on their trust of robots in an emergency scenario. *Paladyn*, 9(1), 137–154. <https://doi.org/10.1515/pjbr-2018-0010>.
- 41 Rossi, S., Conti, D., Garramone, F., Santangelo, G., Staffa, M., Varrasi, S., & Di Nuovo, A. (2020). The role of personality factors and empathy in the acceptance and performance of a social robot for psychometric evaluations. *Robotics*, 9(2), 39. <https://doi.org/10.3390/robotics9020039>.

- 42 Salem, M., Lakatos, G., Amirabdollahian, F., & Dautenhahn, K. (2015). *Would you trust a (faulty) robot? proceedings of the tenth annual ACM/IEEE international conference on Human-robot interaction*, 141–148. <https://doi.org/10.1145/2696454.2696497>.
- 43 Sarkar, S., Araiza-Illan, D., & Eder, K. (2017). *Effects of faults, experience, and personality on trust in a robot co-worker*. ArXiv Preprint, 1703.02335. <http://arxiv.org/abs/1703.02335>.
- 44 Schaefer, K. E., & Scribner, D. R. (2015). Individual differences, trust, and vehicle autonomy. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 59(1), 786–790. <https://doi.org/10.1177/1541931215591242>.
- 45 Schaefer, K. E., & Straub, E. R. (2016). Will passengers trust driverless vehicles? Removing the steering wheel and pedals. *2016 IEEE International Multi-Disciplinary Conference on Cognitive Methods in Situation Awareness and Decision Support (CogSIMA)*, 159–165. <https://doi.org/10.1109/COGSIMA.2016.7497804>.
- 46 Schmidt, P., & Biessmann, F. (2020). Calibrating Human-AI collaboration: impact of risk, ambiguity and transparency on algorithmic bias. In A. H. et Al. (Ed.), *machine learning and knowledge extraction*. CD-MAKE 2020. *Lecture Notes in Computer Science* (Vol. 12279, pp. 431–449). Springer. https://doi.org/10.1007/978-3-030-57321-8_24.
- 47 Schrum, M. L., Neville, G., Johnson, M., Moorman, N., Paleja, R., Feigh, K. M., & Gombolay, M. C. (2021). Effects of social factors and team dynamics on adoption of collaborative robot autonomy. *ACM/IEEE International Conference on Human-Robot Interaction*, 149–157. <https://doi.org/10.1145/3434073.3444649>.
- 48 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>.
- 49 Sorrentino, A., Mancioppi, G., Coviello, L., Cavallo, F., & Fiorini, L. (2021). Feasibility study on the role of personality, emotion, and engagement in socially assistive robotics: a cognitive assessment scenario. *Informatics*, 8(2), 23. <https://doi.org/10.3390/informatics8020023>.
- 50 Szalma, J. L., & Taylor, G. S. (2011). Individual differences in response to automation: The five factor model of personality. *Journal of Experimental Psychology: Applied*, 17(2), 71–96. <https://doi.org/10.1037/a0024170>.
- 51 Tenhundfeld, N. L., de Visser, E. J., Ries, A. J., Finomore, V. S., & Tossell, C. C. (2020). Trust and distrust of automated parking in a tesla model x. *Human Factors*, 62(2), 194–210. <https://doi.org/10.1177/0018720819865412>.
- 52 Tong, S. T., Corriero, E. F., Matheny, R. G., & Hancock, J. T. (2018). Online daters’ willingness to use recommender technology for mate selection decisions. *IntRS Workshop*, 2225, 45–52.
- 53 Voinescu, A., Morgan, P. L., Alford, C., & Caleb-Solly, P. (2018). Investigating older adults’ preferences for functions within a human-machine interface designed for fully autonomous vehicles. In Z. J. & S. G. (Eds.), *human aspects of IT for the aged population. Applications in health, assistance, and entertainment*. ITAP 2018. *Lecture Notes in Computer Science* (Vol. 10927, pp. 445–462). Springer. https://doi.org/10.1007/978-3-319-92037-5_32.
- 54 Xu, K. (2019). First encounter with robot Alpha: How individual differences interact with vocal and kinetic cues in users’ social responses. *New Media & Society*, 21(11–12), 2522–2547. <https://doi.org/10.1177/146144819851479>.
- 55 Yang, F., Huang, Z., Scholtz, J., & Arendt, D. L. (2020). *How do visual explanations foster end users’ appropriate trust in machine learning? 25th International Conference on Intelligent User Interfaces (IUI ’20)*, 189–201. <https://doi.org/10.1145/3377325.3377480>.
- 56 Youn, S., & Jin, S. V. (2021). “In A.I. we trust?” The effects of parasocial interaction and technopion versus luddite ideological views on chatbot-based customer relationship management in the emerging “feeling economy.” *Computers in Human Behavior*, 119, 106721. <https://doi.org/10.1016/j.chb.2021.106721>.
- 57 Zhang, T., Tao, D., Qu, X., Zhang, X., Zeng, J., Zhu, H., & Zhu, H. (2020). Automated vehicle acceptance in China: Social influence and initial trust are key determinants. *Transportation Research Part C: Emerging Technologies*, 112, 220–233. <https://doi.org/10.1016/j.trc.2020.01.027>.
- 58 Zhou, J., Luo, S., & Chen, F. (2020). Effects of personality traits on user trust in human-machine collaborations. *Journal on Multimodal User Interfaces*, 14(4), 387–400. <https://doi.org/10.1007/s12193-020-00329-9>.

Appendix 2

The following table summarizes the major characteristics of the 58 reviewed papers and their underlying studies.

No.	Discipline	Universal traits	Specific traits	Context	Method	Sample size (N)	Mean age (years)	Female (in %)	Country (data collection)
1 P	CompSci	Big Five	Trust propensity	Domestic robot for food preparation	St1: Survey St2: Onl Exp	St1: 173 St2: 138	St1: 36.75 St2: 38.46	St1: 39% St2: n.a.	St1: Canada St2: Canada
2 J	CompSci	Big Five		AI-driven healthcare technologies	Survey	936	37.1	45%	USA
3 J	IS	Big Five		AI-based voice shopping	Survey	224	< 21 (6), 21–40 (149), 41–55 (52), 56–74 (16), > 74 (1)	44%	USA
4 P	HCI	Big Five		AI-enabled user interfaces	Survey	211	15–25 (4%), 26–35 (46%), 36–45 (28%), > 45 (22%)	41%	Different countries
5 P	Ergonom	Big Five		Air traffic control automation	Survey	360	20.92	n.a.	Different countries
6 P	CompSci		Locus of control	Human-robot interaction	Lab Exp	20	30.4	35%	UK
7 J	HCI		Locus of control	Autonomous vehicle	Survey	552	< 30 (31.9%), 30–39 (56.2%), 40–49 (7.6%), 50–59 (2.5%), > 59 (1.8%)	30.1%	Korea
8 P	IS		Computer anxiety, self-efficacy, need for interaction	Self-service kiosks in healthcare	Survey	192	18–25 (26), 26–40 (77), 41–55 (62), 56+ (20), missing (7)	56%	South Africa
9 J	CompSci	Big Five	Need for cognition	Intelligent tutoring system	Lab Exp	47	n.a.	79%	Canada
10 P	CompSci		Locus of control	In-vehicle agents	Onl Exp	100	33	26%	Different countries
11 P	IS	E		Automated recommendation agent	Lab Exp	64	23	52%	USA
12 P	IS	Big Five		Human-robot interaction	Lab study	58	21.69	n.a.	USA
13 P	HCI		Propensity to take risks	Autonomous system	Survey	344	23–38 (54.06%)	48.58%	USA
14 P	Ergonom	Big Five	Dispositional trust in automation, trust propensity	Autonomous system	Survey	344	23–38 (54.06%)	48.58%	USA

No.	Discipline	Universal traits	Specific traits	Context	Method	Sample size (N)	Mean age (years)	Female (in %)	Country (data collection)
15 P	IS		Trust propensity, distrust propensity, suspicion propensity	Interaction with an automated partner in a trust game context	Lab Exp	49	23.04	57%	USA
16 J	HCI		Attachment style	Artificial intelligence	St1: Survey St2+3: Onl Exp	St1: 248 St2: 374 St3: 272	St1: 18–80 St2: 18–78 St3: 18–82	St1: 72% St2: 61% St3: 65%	St1: USA St2: USA St3: USA
17 P	IS		Need for interaction, self-efficacy, technological innovativeness, novelty seeking	Intelligent Personal Assistants	Survey	168	32.7	46.4%	Germany
18 P	CompSci	A, E		Intelligent Virtual Agents	Lab Exp	55	22.56	n.a.	Australia
19 P	CompSci	Eysenck		Human-robot interaction (trust game)	Lab Exp	55	22.6	67%	Japan
20 P	Ergonom	O, N	Need for cognition, locus of control, trust propensity	Human-robot interaction	Lab Exp	28	18–55	61%	USA
21 J	HCI		Innovativeness, driving enjoyment	Autonomous vehicles	Survey	369	31	56%	Germany
22 P	HCI		Trust propensity, dispositional trust in automation	Automated security officer	Onl Exp	156	< 30 (43.6%), 30–60 (50.7%), > 60 (5.8%)	38%	Different countries
23 P	CompSci	A, E, N	Dispositional trust in automation, trust propensity	Automated agent (airport)	Onl Exp	156	< 30 (43.6%), 30–60 (50.7%), > 60 (5.8%)	38%	Different countries
24 P	CompSci		Immersive tendency, need to belong,	Social robots	Lab Exp	20	22.75	60%	USA
25 J	Ergonom		Perfect Automation Schema (PAS)	Autonomous security robot	Onl Exp	233	33.75	46%	n.a.
26 J	Ergonom	Big Five	Self-esteem, locus of control, affinity for technology, self-efficacy, trust propensity, dispositional trust in automation	Advanced driver-assistance system	Survey	St1: 274 St2: 149	St1: 33.88 St2: 29.96	St1: 58% St2: 62%	St1: Germany St2: Germany
27 J	Psychol		Depressiveness, self-efficacy, self-esteem, locus of control	Automated driving system	Lab study	47	27.45	57%	Germany
28 J	Ergonom		Perfect Automation Schema (PAS)	St1+2: Human-robot interaction St3: Advanced automated safety system	St1: Lab Exp St2: Lab study St3: Survey	St1: 38 St2: 130 St3: 100	St1: 35 St2: 24 St3: n.a.	St1: 39% St2: 67% St3: n.a.	St1: USA St2: USA St3: USA

No.	Discipline	Universal traits	Specific traits	Context	Method	Sample size (N)	Mean age (years)	Female (in %)	Country (data collection)
29 P	CompSci	Big Five	Perfect Automation Schema (PAS)	Autonomous security robot	Survey	316	39	n.a.	n.a.
30 P	CompSci		Belief in paranormal entities	Human-robot interaction	Survey	59	44.6	24%	Japan
31 J	CompSci		Attitude towards robots	Human-robot interaction	Onl Exp	82	18–40	35%	n.a.
32 J	IS	Myers-Briggs		Automated decision aids (nursing)	Lab study	55	35	100%	USA
33 J	Ergonom	E	Dispositional trust in automation	Automatic weapon detector	Lab Exp	255	19.25	n.a.	USA
34 J	Ergonom		Implicit attitude towards automated systems, dispositional trust in automation	Automated baggage inspector	Onl Exp	69	25	77%	USA
35 J	Psychol		Dispositional trust in automation, attitude towards robots	Home assistance robot	Lab Exp	28	30.32	57%	Germany
36 P	IS	HEXACO		Chatbot	Survey	112	21–30 (44.6%), 31–40 (27.7%), 41–50 (12.5%), 51–60 (8%), > 60 (7.2%)	51.8%	Different countries
37 J	Psychol	Big Five	Robot use self-efficacy	Robot and AI support in decision-making trust game	Onl Exp	1,077	37.39	50.6%	USA
38 P	Ergonom		Mental model	Navigation technology	Lab Exp	27	35.67	41%	USA
39 J	Ergonom		Expectancy that automation is trustworthy	Baggage screening technology	Lab Exp	225	19.94	48%	USA
40 J	CompSci	Big Five	Trust propensity	Home companion robot	Onl Exp	200	33.56	43%	Different countries
41 J	CompSci	NEO-PI-3		Socially assistive robot	Lab Exp	19	61	42%	Italy
42 P	CompSci	E, N		Human-robot interaction	Lab Exp	40	37.95	55%	UK
43 *	arXiv	Big Five	Attitude towards robots	Human-robot interaction	Lab Exp	18	18–24 (12), 25–29 (4), 30–39 (2)	22%	UK
44 P	Ergonom		Dislike of driving, fatigue proneness, thrill seeking, driving aggression	Autonomous passenger vehicle	Lab Exp	24	n.a.	46%	USA
45 P	Ergonom		Intellect (having a vivid imagination)	Driverless vehicle	Lab Exp	20	18–64	30%	USA
46 P	CompSci		Affinity to risk	Machine learning support in decision-making	Onl Exp	248	n.a.	n.a.	n.a.

No.	Discipline	Universal traits	Specific traits	Context	Method	Sample size (N)	Mean age (years)	Female (in %)	Country (data collection)
47 P	CompSci		Propensity to trust	Human-robot interaction	Lab Exp	42	21.1	52.4%	USA
48 J	Other	Big Five	Locus of control	AI support in decision-making card game	Onl Exp	171	22.6	78%	Different countries
49 J	CompSci	Big Five		Socially assistive robotics	Lab study	8	83.33	38%	Italy
50 J	Psychol	Big Five		Simulated uninhabited ground vehicle	Lab Exp	161	19.8	32%	USA
51 J	Ergonom		Risk-taking behavior, self-confidence	Partially automated parking with a Tesla Model X	Field study	23	18–22	17%	USA
52 P	HCI		Need for cognition	Recommender technology	Lab Exp	129	20.35	76%	USA
53 P	HCI		Impulsive sensation seeking, aggression hostility, sociability, activity, neuroticism anxiety, attitude towards computers	Autonomous Vehicle	Lab study	31	67.52	61.3%	UK
54 J	Other		Attitude towards robots	Social robot	Lab Exp	110	20.44	50%	USA
55 P	HCI		Propensity to trust	Visual explanations for machine learning classifiers	Lab Exp	33	25–60	58%	USA
56 J	HCI		Ideological view (technopian vs. luddite)	Chatbot	Onl Exp	602	40.8	43%	USA
57 J	Other	Big Five	Sensation seeking	Automated Vehicle	Survey	604	31.38	42.9%	China
58 J	HCI	Big Five		Water pipe failure prediction technology	Lab Exp	42	30.4	24%	Australia

Notes: We harmonized terminology across the 58 papers in the column “Specific traits”. In particular, we use the term “trust propensity” despite the fact that similar terms like “interpersonal propensity to trust”, “disposition to trust”, “dispositional trust”, “dispositional interpersonal trust”, “dispositional trust in humans”, or “trust disposition” are used in the literature. Abbreviations: *J* Journal, *P* Proceedings, * = arXiv. *CompSci* Computer Science, Informatics, Robotics, *HCI* Human-Computer Interaction, *Ergonom* Ergonomics, Human Factors, *Psychol* Psychology. *IS* Information Systems, *A* Agreeableness, *E* Extraversion, *N* Neuroticism, *O* Openness, *Lab Exp* Laboratory Experiment, *Onl Exp* Online Experiment, *St* Study. *n.a.* not available

Funding Open access funding provided by University of Applied Sciences Upper Austria.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Adam, M. T. P., Gimpel, H., Maedche, A., & Riedl, R. (2017). Design blueprint for stress-sensitive adaptive Enterprise systems. *Business & Information Systems Engineering*, 59(4), 277–291. <https://doi.org/10.1007/s12599-016-0451-3>
- Adorno, T., Frenkel-Brunswick, E., Levinson, D., & Sanford, N. (1950). *The authoritarian personality*. Harper.
- Akash, K., Hu, W.-L., Jain, N., & Reid, T. (2018). A classification model for sensing human trust in machines using EEG and GSR. *ACM Transactions on Interactive Intelligent Systems*, 8(4), 1–20. <https://doi.org/10.1145/3132743>
- Aliasghari, P., Ghafurian, M., Nehaniv, C. L., & Dautenhahn, K. (2021). Effect of domestic trainee robots' errors on human teachers' trust. In *2021 30th IEEE International Conference on Robot and Human Interactive Communication, RO-MAN 2021* (pp. 81–88). <https://doi.org/10.1109/RO-MAN50785.2021.9515510>

- Antes, A. L., Burrous, S., Sisk, B. A., Schuelke, M. J., Keune, J. D., & DuBois, J. M. (2021). Exploring perceptions of healthcare technologies enabled by artificial intelligence: an online, scenario-based survey. *BMC Medical Informatics and Decision Making*, 21(1), 221. <https://doi.org/10.1186/s12911-021-01586-8>
- Ashton, M. C., & Lee, K. (2016). Age trends in HEXACO-PI-R self-reports. *Journal of Research in Personality*, 64, 102–111. <https://doi.org/10.1016/j.jrp.2016.08.008>
- Ashton, M. C., Lee, K., Perugini, M., Szarota, P., de Vries, R. E., Di Blas, L., Boies, K., & De Raad, B. (2004). A six-factor structure of personality-descriptive adjectives: Solutions from psycholexical studies in seven languages. *Journal of Personality and Social Psychology*, 86(2), 356–366. <https://doi.org/10.1037/0022-3514.86.2.356>
- Astor, P. J., Adam, M. T. P., Jerčić, P., Schaaff, K., & Weinhardt, C. (2013). Integrating biosignals into information systems: A NeuroIS tool for improving emotion regulation. *Journal of Management Information Systems*, 30(3), 247–278. <https://doi.org/10.2753/MIS0742-1222300309>
- 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>
- Bawack, R. E., Wamba, S. F., & Carillo, K. D. A. (2021). Exploring the role of personality, trust, and privacy in customer experience performance during voice shopping: Evidence from SEM and fuzzy set qualitative comparative analysis. *International Journal of Information Management*, 58, 102309. <https://doi.org/10.1016/j.ijinfomgt.2021.102309>
- Berente, N., Gu, B., Recker, J., & Santhanam, R. (2021). Managing Artificial Intelligence. *MIS Quarterly*, 45(3), 1433–1450. <https://doi.org/10.25300/MISQ/2021/16274>
- Böckle, M., Yeboah-Antwi, K., & Kouris, I. (2021). Can you trust the black box? the effect of personality traits on trust in AI-enabled user interfaces. In Degen, H., & Ntoa, S. (Ed.), *Artificial Intelligence in HCI. HCII 2021. Lecture notes in computer science* (Vol. 12797, pp. 3–20). Springer. https://doi.org/10.1007/978-3-030-77772-2_1
- Byrne, E. A., & Parasuraman, R. (1996). Psychophysiology and adaptive automation. *Biological Psychology*, 42(3), 249–268. [https://doi.org/10.1016/0301-0511\(95\)05161-9](https://doi.org/10.1016/0301-0511(95)05161-9)
- Cacioppo, J. T., Petty, R. E., & Feng Kao, C. (1984). The efficient assessment of need for cognition. *Journal of Personality Assessment*, 48(3), 306–307. https://doi.org/10.1207/s15327752jpa4803_13
- Cattell, R. B., Eber, H. W., & Tatsuoaka, M. M. (1970). *The handbook for the sixteen personality factor questionnaire*. Edited by the Institute for Personality and Ability Testing.
- Chen, W., & Hirschheim, R. (2004). A paradigmatic and methodological examination of information systems research from 1991 to 2001. *Information Systems Journal*, 14(3), 197–235. <https://doi.org/10.1111/j.1365-2575.2004.00173.x>
- 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/10.1177/15419312163601192>
- Chiou, M., McCabe, F., Grigoriou, M., & Stolk, R. (2021). Trust, shared understanding and locus of control in mixed-initiative robotic systems. In *2021 30th IEEE International Conference on Robot & Human Interactive Communication (RO-MAN)* (pp. 684–691). <https://doi.org/10.1109/RO-MAN50785.2021.9515476>
- Choi, J. K., & Ji, Y. G. (2015). Investigating the importance of trust on adopting an autonomous vehicle. *International Journal of Human-Computer Interaction*, 31(10), 692–702. <https://doi.org/10.1080/10447318.2015.1070549>
- Cohen, J. F., & Sergay, S. D. (2011). An empirical study of health consumer beliefs, attitude and intentions toward the use of self-service kiosks. In *17th Americas Conference on Information Systems 2011, AMCIS 2011 Proceedings - All Submissions* (Vol. 46, pp. 403–412).
- Collins, C., Dennehy, D., Conboy, K., & Mikalef, P. (2021). Artificial intelligence in information systems research: A systematic literature review and research agenda. *International Journal of Information Management*, 60, 102383. <https://doi.org/10.1016/j.ijinfomgt.2021.102383>
- Conati, C., Barral, O., Putnam, V., & Rieger, L. (2021). Toward personalized XAI: A case study in intelligent tutoring systems. *Artificial Intelligence*, 298, 103503. <https://doi.org/10.1016/j.artint.2021.103503>
- Costa Jr., P. T., & McCrae, R. R. (1992). *Revised NEO personality inventory (NEO-PI-R) and NEO five-factor inventory (NEO-FFI) manual*. Psychological Assessment Resources.
- Cramer, H., Evers, V., Kemper, N., & Wielinga, B. (2008). Effects of autonomy, traffic conditions and driver personality traits on attitudes and trust towards in-vehicle agents. *IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology*, 2008, 477–482. <https://doi.org/10.1109/WIAT.2008.326>
- Dabholkar, P. A. (1992). The role of prior behavior and category-based affect in on-site service encounters. In J. F. Sherry & B. Sternthal (Eds.), *Diversity in consumer behavior: Vol. XIX* (pp. 563–569). Association for Consumer Research.
- Demazure, T., Karran, A., Léger, P.-M., Labonté-LeMoyne, É., Sénécal, S., Fredette, M., & Babin, G. (2021). Enhancing sustained attention. *Business & Information Systems Engineering*, 63(6), 653–668. <https://doi.org/10.1007/s12599-021-00701-3>
- Devaraj, U. S., Easley, R. F., & Michael Crant, J. (2008). How does personality matter? Relating the five-factor model to technology acceptance and use. *Information Systems Research*, 19(1), 93–105. <https://doi.org/10.1287/isre.1070.0153>
- Dimoka, A., Banker, R. D., Benbasat, I., Davis, F. D., Dennis, A. R., Gefen, D., Gupta, A., Ischebeck, A., Henning, P. H., Pavlou, P. A., Müller-Putz, G., Riedl, R., vom Brocke, J., & Weber, B. (2012). On the use of neurophysiological tools in IS research: Developing a research agenda for NeuroIS. *MIS Quarterly*, 36(3), 679–702. <https://doi.org/10.2307/41703475>
- Dwivedi, Y. K., Hughes, L., Ismagilova, E., Aarts, G., Coombs, C., Crick, T., Duan, Y., Dwivedi, R., Edwards, J., Eirug, A., Galanos, V., Ilavarasan, P. V., Janssen, M., Jones, P., Kar, A. K., Kizgin, H., Kronemann, B., Lal, B., Lucini, B., Medaglia, R., Le Meunier-FitzHugh, K., Le Meunier-Fitz Hugh, L. C., Misra, S., Mogaji, E., Kumar Sharma, S., Bahadur Singh, J., Raghavan, V., Ramanu, R., Rana, N. P., Samothrakis, S., Spencer, J., Tamilmani, K., Tubadji, A., Walton, P., & Williams, M. D. (2021). Artificial intelligence (AI): Multidisciplinary perspectives on emerging challenges, opportunities, and agenda for research, practice and policy. *International Journal of Information Management*, 57(101994), 1–47. <https://doi.org/10.1016/j.ijinfomgt.2019.08.002>
- Dzindolet, M. T., Pierce, L. G., Beck, H. P., & Dawe, L. A. (2002). The perceived utility of human and automated aids in a visual detection task. *Human Factors*, 44(1), 79–94. <https://doi.org/10.1518/0018720024494856>
- Elson, J. S., Derrick, D., & Ligon, G. (2018). Examining trust and reliance in collaborations between humans and automated agents. In *Proceedings of the 51st Hawaii International Conference on System Sciences* (pp. 430–439). <https://doi.org/10.24251/HICSS.2018.056>
- Elson, J. S., Derrick, D., & Ligon, G. (2020). Trusting a humanoid robot: exploring personality and trusting effects in a human-robot partnership. In *Proceedings of the 53rd Hawaii International*

- Conference on System Sciences* (pp. 543–552). <https://doi.org/10.24251/HICSS.2020.067>
- Eysenck, H. J. (1947). *Dimensions of personality*. Kegan Paul.
- Eysenck, H. J., & Eysenck, S. B. G. (1976). *Psychoticism as a dimension of personality*. Hodder & Stoughton.
- Ferronato, P., & Bashir, M. (2020a). An examination of dispositional trust in human and autonomous system interactions. In K. M. (Ed.), *human-computer interaction. human values and quality of life. HCII 2020. Lecture notes in computer science* (Vol. 12183, pp. 420–435). Springer. https://doi.org/10.1007/978-3-030-49065-2_30
- Ferronato, P., & Bashir, M. (2020b). Does the propensity to take risks influence human interactions with autonomous systems? In I. Corradini, E. Nardelli, & T. Ahram (Eds.), *Advances in human factors in cybersecurity. AHFE 2020. Advances in intelligent systems and computing* (Vol. 1219, pp. 23–29). Springer. https://doi.org/10.1007/978-3-030-52581-1_4
- Funder, D. C. (2001). Personality. *Annual Review of Psychology*, 52(1), 197–221. <https://doi.org/10.1146/annurev.psych.52.1.197>
- Gefen, D. (2000). E-commerce: The role of familiarity and trust. *Omega*, 28(6), 725–737. [https://doi.org/10.1016/S0305-0483\(00\)00021-9](https://doi.org/10.1016/S0305-0483(00)00021-9)
- Gibson, A. M., Alarcon, G. M., Jessup, S. A., & Capiola, A. (2020). “Do you still trust me?” effects of personality on changes in trust during an experimental task with a human or robot partner. In *Proceedings of the 53rd Hawaii International Conference on System Sciences* (pp. 5099–5108).
- Gillath, O., Ai, T., Branicky, M. S., Keshmiri, S., Davison, R. B., & Spaulding, R. (2021). Attachment and trust in artificial intelligence. *Computers in Human Behavior*, 115, 106607. <https://doi.org/10.1016/j.chb.2020.106607>
- Gist, M. E. (1987). Self-efficacy: Implications for organizational behavior and human resource management. *The Academy of Management Review*, 12(3), 472. <https://doi.org/10.2307/258514>
- 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>
- Goldberg, L. R. (1990). An alternative “description of personality”: The Big-Five factor structure. *Journal of Personality and Social Psychology*, 59(6), 1216–1229. <https://doi.org/10.1037/0022-3514.59.6.1216>
- Goldberg, L. R. (1993). The structure of phenotypic personality traits. *American Psychologist*, 48(1), 26–34. <https://doi.org/10.1037/0003-066X.48.1.26>
- Hampson, S. E., Goldberg, L. R., & John, O. P. (1987). Category-breadth and social-desirability values for 573 personality terms. *European Journal of Personality*, 1(4), 241–258. <https://doi.org/10.1002/per.2410010405>
- 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>
- Handrich, M. (2021). Alexa, you freak me out – Identifying drivers of innovation resistance and adoption of Intelligent Personal Assistants. *ICIS 2021 Proceedings*, 11, 1–17.
- Hanna, N., & Richards, D. (2015). The influence of users’ personality on the perception of intelligent virtual agents’ personality and the trust within a collaborative context. In F. K. et al. (Eds.), *CARE-MFSC 2015* (Vol. 541, pp. 31–47). CCIS, Springer. https://doi.org/10.1007/978-3-319-24804-2_3
- Haring, K. S., Matsumoto, Y., & Watanabe, K. (2013). How do people perceive and trust a lifelike robot. *Proceedings of the World Congress on Engineering and Computer Science*, 2013(1), 425–430.
- Harriott, C. E., Garver, S., & Cunha, M. (2018). A motivation for co-adaptive human-robot interaction. In C. J. (Ed.), *Advances in human factors in robots and unmanned systems, advances in intelligent systems and computing* (Vol. 595, pp. 148–160). Springer. https://doi.org/10.1007/978-3-319-60384-1_15
- Harriott, C. E., Garver, S., & Cunha, M. (2018). A motivation for co-adaptive human-robot interaction. In J. Chen (Ed.), *Advances in human factors in robots and unmanned systems. AHFE 2017. Advances in intelligent systems and computing* (Vol. 595). Springer. https://doi.org/10.1007/978-3-319-60384-1_15
- Harter, S. (1993). Causes and consequences of low self-esteem in children and adolescents. In R. Baumeister (Ed.), *Self-esteem: The puzzle of low self-regard* (pp. 87–116). Plenum. https://doi.org/10.1007/978-1-4684-8956-9_5
- Hegner, S. M., Beldad, A. D., & Brunswick, G. J. (2019). In automatic we trust: Investigating the impact of trust, control, personality characteristics, and extrinsic and intrinsic motivations on the acceptance of autonomous vehicles. *International Journal of Human-Computer Interaction*, 35(19), 1769–1780. <https://doi.org/10.1080/10447318.2019.1572353>
- Hess, T., Fuller, M., & Campbell, D. (2009). Designing interfaces with social presence: Using vividness and extraversion to create social recommendation agents. *Journal of the Association for Information Systems*, 10(12), 889–919. <https://doi.org/10.17705/ljais.00216>
- Hevner, A. R., March, S. T., Park, J., & Ram, S. (2004). Design science in information systems research. *MIS Quarterly*, 28(1), 75–105. <https://doi.org/10.2307/25148625>
- Hoff, K. A., & Bashir, M. (2015). Trust in automation: Integrating empirical evidence on factors that influence trust. *Human Factors*, 57(3), 407–434. <https://doi.org/10.1177/0018720814547570>
- Hofstede, G., & McCrae, R. R. (2004). Personality and culture revisited: Linking traits and dimensions of culture. *Cross-Cultural Research*, 38(1), 52–88. <https://doi.org/10.1177/1069397103259443>
- Huang, H.-Y., & Bashir, M. (2017). Personal influences on dynamic trust formation in human-agent interaction. In *Proceedings of the 5th International Conference on Human Agent Interaction* (pp. 233–243). <https://doi.org/10.1145/3125739.3125749>
- Huang, H.-Y., Twidale, M., & Bashir, M. (2020). “If you agree with me, do I trust you?”: An examination of human-agent trust from a psychological perspective. In Y. Bi & Y. Bi (Eds.), *Intelligent Systems and Applications. IntelliSys 2019. Advances in Intelligent Systems and Computing* (Vol. 1038, pp. 994–1013). Springer. https://doi.org/10.1007/978-3-030-29513-4_73
- Jacovi, A., Marasovi, A., Miller, T., & Goldberg, Y. (2021). Formalizing trust in artificial intelligence: Prerequisites, causes and goals of human trust in AI. *Proceedings of the 2021 ACM conference on fairness, accountability, and transparency*, pp. 624–635. <https://doi.org/10.1145/3442188.3445923>
- Jiang, J. J., Klein, G., & Carr, C. L. (2002). Measuring information system service quality: SERVQUAL from the other side. *MIS Quarterly*, 26(2), 145–166. <https://www.jstor.org/stable/4132324>
- John, O. P., & Srivastava, S. (1999). The Big-Five trait taxonomy: History, measurement, and theoretical perspectives. In L. Pervin & O. P. John (Eds.), *Handbook of personality: Theory and research*. The Guilford Press.
- John, O. P., Naumann, L. P., & Soto, C. J. (2008). Paradigm shift to the integrative Big Five trait taxonomy: History, measurement, and conceptual issues. In O. P. John, R. W. Robins, & L. A. Pervin (Eds.), *Handbook of personality: Theory and research* (3rd ed., pp. 114–158). The Guilford Press.
- Jung, C. G. (1923). *Psychological types*. Harcourt, Brace.
- Kampman, O., Siddique, F. B., Yang, Y., & Fung, P. (2019). Adapting a virtual agent to user personality. In M. Eskenazi, L. Devillers, & J. Mariani (Eds.), *Advanced social interaction with agents. Lecture notes in electrical engineering* (Vol. 510, pp. 111–118). Springer. https://doi.org/10.1007/978-3-319-92108-2_13

- Kim, K. J., Park, E., Sundar, S. S., & del Pobil, A. P. (2012). The effects of immersive tendency and need to belong on human-robot interaction. In *Proceedings of the Seventh Annual ACM/IEEE International Conference on Human-Robot Interaction - HRI '12* (pp. 207–208). <https://doi.org/10.1145/2157689.2157758>
- Kim, W., Kim, N., Lyons, J. B., & Nam, C. S. (2020). Factors affecting trust in high-vulnerability human-robot interaction contexts: A structural equation modelling approach. *Applied Ergonomics*, *85*, 103056. <https://doi.org/10.1016/j.apergo.2020.103056>
- Klein, H. K., & Hirschheim, R. (2008). The structure of the IS discipline reconsidered: Implications and reflections from a community of practice perspective. *Information and Organization*, *18*(4), 280–302. <https://doi.org/10.1016/j.infoandorg.2008.05.001>
- Kraus, J., Scholz, D., & Baumann, M. (2020a). What's driving me? Exploration and validation of a hierarchical personality model for trust in automated driving. *Human Factors*, *63*(6), 1076–1105. <https://doi.org/10.1177/0018720820922653>
- Kraus, J., Scholz, D., Messner, E.-M., Messner, M., & Baumann, M. (2020b). 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>
- Lankton, N., McKnight, D. H., & Tripp, J. (2015). Technology, humanness, and trust: Rethinking trust in technology. *Journal of the Association for Information Systems*, *16*(10), 880–918. <https://doi.org/10.17705/1jais.00411>
- Lee, J., & Moray, N. (1992). Trust, control strategies and allocation of function in human-machine systems. *Ergonomics*, *35*(10), 1243–1270. <https://doi.org/10.1080/00140139208967392>
- 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
- Leung, A. K.-Y., & Cohen, D. (2011). Within- and between-culture variation: Individual differences and the cultural logics of honor, face, and dignity cultures. *Journal of Personality and Social Psychology*, *100*(3), 507–526. <https://doi.org/10.1037/a0022151>
- Lyons, J. B., & Guznov, S. Y. (2019). Individual differences in human-machine trust: A multi-study look at the perfect automation schema. *Theoretical Issues in Ergonomics Science*, *20*(4), 440–458. <https://doi.org/10.1080/1463922X.2018.1491071>
- Lyons, J. B., Nam, C. S., Jessup, S. A., Vo, T. Q., & Wynne, K. T. (2020). The role of individual differences as predictors of trust in autonomous security robots. *IEEE International Conference on Human-Machine Systems (ICHMS)*, (pp. 1–5). <https://doi.org/10.1109/ICHMS49158.2020.9209544>
- Maier, C. (2012). Personality within information systems research: A literature analysis. *Proceedings of the European Conference on Information Systems*, 101.
- March, S. T., & Smith, G. F. (1995). Design and natural science research on information technology. *Decision Support Systems*, *15*(4), 251–266. [https://doi.org/10.1016/0167-9236\(94\)00041-2](https://doi.org/10.1016/0167-9236(94)00041-2)
- Mason, R. O., McKenney, J. L., & Copeland, D. G. (1997). An historical method for MIS research: Steps and assumptions. *MIS Quarterly*, *21*(3), 307. <https://doi.org/10.2307/249499>
- Matsui, T. (2021). Relationship between users' trust in robots and belief in paranormal entities. In *Proceedings of the 9th International Conference on Human-Agent Interaction* (pp. 252–256). <https://doi.org/10.1145/3472307.3484666>
- Matthews, G., Lin, J., Panganiban, A. R., & Long, M. D. (2020). Individual differences in trust in autonomous robots: Implications for transparency. *IEEE Transactions on Human-Machine Systems*, *50*(3), 234–244. <https://doi.org/10.1109/THMS.2019.2947592>
- Matthews, G., Hancock, P. A., Lin, J., Panganiban, A. R., Reinerman-Jones, L. E., Szalma, J. L., & Wohleber, R. W. (2021). Evolution and revolution: Personality research for the coming world of robots, artificial intelligence, and autonomous systems. *Personality and Individual Differences*, *169*, 109969. <https://doi.org/10.1016/j.paid.2020.109969>
- Mayer, R. C., Davis, J. H., & Schoorman, F. D. (1995). An integrative model of organizational trust. *Academy of Management Review*, *20*(3), 709–734. <https://doi.org/10.5465/amr.1995.9508080335>
- McBride, M., Carter, L., & Ntuen, C. (2012). The impact of personality on nurses' bias towards automated decision aid acceptance. *International Journal of Information Systems and Change Management*, *6*(2), 132–146. <https://doi.org/10.1504/IJISCM.2012.051148>
- McCarthy, J. L., Minsky, M. L., Rochester, N., & Shannon, C. E. (1955). *A proposal for the Dartmouth summer research project on artificial intelligence*. <http://jmc.stanford.edu/articles/dartmouth/dartmouth.pdf>
- McCrae, R. R. (2000). Trait psychology and the revival of personality and culture studies. *American Behavioral Scientist*, *44*(1), 10–31. <https://doi.org/10.1177/00027640021956062>
- McCrae, R. R. (2004). Human nature and culture: A trait perspective. *Journal of Research in Personality*, *38*(1), 3–14. <https://doi.org/10.1016/j.jrp.2003.09.009>
- McCrae, R. R., & Costa, P. T. J. (1987). Validation of the five-factor model of personality across instruments and observers. *Journal of Personality and Social Psychology*, *52*(1), 81–90. <https://doi.org/10.1037/0022-3514.52.1.81>
- McCrae, R. R., & Costa, P. T. J. (1997). Personality trait structure as a human universal. *American Psychologist*, *52*(5), 509–516. <https://doi.org/10.1037/0003-066X.52.5.509>
- McCrae, R. R., & Costa, P. T. J. (1999). A Five-Factor Theory of personality. In A. Pervin & O. P. John (Eds.), *Handbook of personality: Theory and research* (2nd ed., pp. 139–153). The Guilford Press.
- McCrae, R. R., Costa, P. T. J., & Martin, T. A. (2005). The NEO-PI-3: A more readable revised NEO personality inventory. *Journal of Personality Assessment*, *84*(3), 261–270. https://doi.org/10.1207/s15327752jpa8403_05
- McElroy, J. C., Hendrickson, A. R., Townsend, A. M., & DeMarie, S. M. (2007). Dispositional factors in internet use: Personality versus cognitive style. *MIS Quarterly*, *31*(4), 809–820. <https://doi.org/10.2307/25148821>
- McKnight, D. H., Cummings, L. L., & Chervany, N. L. (1998). Initial trust formation in new organizational relationships. *Academy of Management Review*, *23*(3), 473–490. <https://doi.org/10.5465/amr.1998.926622>
- 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>
- 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>
- Merritt, S. M., Unnerstall, J. L., Lee, D., & Huber, K. (2015). Measuring individual differences in the perfect automation schema. *Human Factors*, *57*(5), 740–753. <https://doi.org/10.1177/0018720815581247>
- Miller, L., Kraus, J., Babel, F., & Baumann, M. (2021). More than a feeling—interrelation of trust layers in human-robot interaction and the role of user dispositions and state anxiety. *Frontiers in Psychology*, *12*, 592711. <https://doi.org/10.3389/fpsyg.2021.592711>
- Montag, C., & Panksepp, J. (2017). Primary emotional systems and personality: An evolutionary perspective. *Frontiers in Psychology*, *8*, 464. <https://doi.org/10.3389/fpsyg.2017.00464>
- Montag, C., Hahn, E., Reuter, M., Spinath, F. M., Davis, K., & Panksepp, J. (2016). The role of nature and nurture for individual

- differences in primary emotional systems: Evidence from a twin study. *PLoS One*, 11(3), e0151405. <https://doi.org/10.1371/journal.pone.0151405>
- Montag, C., Elhai, J. D., & Davis, K. L. (2021). A comprehensive review of studies using the affective neuroscience personality scales in the psychological and psychiatric sciences. *Neuroscience & Biobehavioral Reviews*, 125, 160–167. <https://doi.org/10.1016/j.neubiorev.2021.02.019>
- Mooradian, T., Renzl, B., & Matzler, K. (2006). Who trusts? Personality, trust and knowledge sharing. *Management Learning*, 37(4), 523–540. <https://doi.org/10.1177/1350507606073424>
- Mount, M. K., Barrick, M. R., Scullen, S. M., & Rounds, J. (2005). Higher-order dimensions of the Big Five personality traits and the big six vocational interest types. *Personnel Psychology*, 58(2), 447–478. <https://doi.org/10.1111/j.1744-6570.2005.00468.x>
- Mühl, K., Strauch, C., Grabmaier, C., Reithinger, S., Huckauf, A., & Baumann, M. (2020). Get ready for being chauffeured: Passenger's preferences and trust while being driven by human and automation. *Human Factors*, 62(8), 1322–1338. <https://doi.org/10.1177/0018720819872893>
- Myers, I. B., McCaulley, M. H., Quenk, N. L., & Hammer, A. L. (1998). *The MBTI® manual: A guide to the development and use of the Myers-Briggs type indicator*. Consulting Psychologists Press.
- Müller, L., Matke, J., Maier, C., Weitzel, T., & Graser, H. (2019). Chatbot acceptance: A latent profile analysis on individuals' trust in conversational agents. In *SIGMIS-CPR 2019 - Proceedings of the 2019 Computers and People Research Conference* (pp. 35–42). <https://doi.org/10.1145/3322385.3322392>
- Nahum-Shani, I., Smith, S. N., Spring, B. J., Collings, L. M., Witkiewitz, K., Tewari, A., & Murphy, S. A. (2018). Just-in-time adaptive interventions (JITAs) in mobile health: Key components and design principles for ongoing health behavior support. *Annals of Behavioral Medicine*, 52(6), 446–462. <https://doi.org/10.1007/s12160-016-9830-8>
- Nam, T. (2019). Citizen attitudes about job replacement by robotic automation. *Futures*, 109, 39–49. <https://doi.org/10.1016/j.futures.2019.04.005>
- Nilsson, N. J. (2010). *The quest for artificial intelligence: A history of ideas and achievements*. Cambridge University Press.
- Orri, M., Pingault, J.-B., Rouquette, A., Lalanne, C., Falissard, B., Herba, C., Côté, S. M., & Berthoz, S. (2017). Identifying affective personality profiles: A latent profile analysis of the affective neuroscience personality scales. *Scientific Reports*, 7(1), 4548. <https://doi.org/10.1038/s41598-017-04738-x>
- 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>
- Panganiban, A. R., Matthews, G., Lin, J., & Long, M. D. (2020). Trust your robot! Individual differences in confidence in robot threat evaluations. Abstracts from the International Society for the Study of individual differences conference 2019. *Personality and Individual Differences*, 157(109684), 36.
- Paravastu, N., Gefen, D., & Creason, S. B. (2014). Understanding trust in IT artifacts. *ACM SIGMIS Database: The DATABASE for Advances in Information Systems*, 45(4), 30–50. <https://doi.org/10.1145/2691517.2691520>
- Peffers, K., Tuunanen, T., Rothenberger, M. A., & Chatterjee, S. (2008). A design science research methodology for information systems research. *Journal of Management Information Systems*, 24(3), 45–77. <https://doi.org/10.2753/MIS0742-1222240302>
- Perelman, B. S., Evans, A. W., & Schaefer, K. E. (2017). Mental model consensus and shifts during navigation system-assisted route planning. *Proceedings of the Human Factors and Ergonomics Society 2017 Annual Meeting*, 61(1), 1183–1187. <https://doi.org/10.1177/1541931213601779>
- Picard, R. W. (1997). *Affective computing*. MIT Press.
- Picard, R. W. (2003). Affective computing: Challenges. *International Journal of Human-Computer Studies*, 59(1–2), 55–64. [https://doi.org/10.1016/S1071-5819\(03\)00052-1](https://doi.org/10.1016/S1071-5819(03)00052-1)
- Pop, V. L., Shrewsbury, A., & Durso, F. T. (2015). Individual differences in the calibration of trust in automation. *Human Factors*, 57(4), 545–556. <https://doi.org/10.1177/0018720814564422>
- Poria, S., Cambria, E., Bajpai, R., & Hussain, A. (2017). A review of affective computing: From unimodal analysis to multimodal fusion. *Information Fusion*, 37, 98–125. <https://doi.org/10.1016/j.inffus.2017.02.003>
- Reuter, M., Panksepp, J., Davis, K. L., & Montag, C. (2017). *Affective neuroscience personality scales (ANPS) – Deutsche Version*. Hogrefe.
- Riedl, R. (2021). Trust and Digitalization: Review of Behavioral and Neuroscience Evidence. In F. Krueger (Ed.), *The Neurobiology of Trust* (pp. 54–76). Cambridge University Press. <https://doi.org/10.1017/9781108770880.005>
- Riedl, R., & Léger, P.-M. (2016). *Fundamentals of NeuroIS – Information systems and the brain*. Springer.
- Rossi, A., Dautenhahn, K., Koay, K. L., & Walters, M. L. (2018). The impact of peoples' personal dispositions and personalities on their trust of robots in an emergency scenario. *Paladyn*, 9(1), 137–154. <https://doi.org/10.1515/pjbr-2018-0010>
- Rossi, S., Conti, D., Garramone, F., Santangelo, G., Staffa, M., Varrasi, S., & Di Nuovo, A. (2020). The role of personality factors and empathy in the acceptance and performance of a social robot for psychometric evaluations. *Robotics*, 9(2), 39. <https://doi.org/10.3390/robotics9020039>
- 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. (1990). Internal versus external control of reinforcement: A case history of a variable. *American Psychologist*, 45(4), 489–493. <https://doi.org/10.1037/0003-066X.45.4.489>
- Rousseau, D. M., Sitkin, S. B., Burt, R. S., & Camerer, C. (1998). Not so different after all: A cross-discipline view of trust. *Academy of Management Review*, 23(3), 393–404. <https://doi.org/10.5465/amr.1998.926617>
- Salem, M., Lakatos, G., Amirabdollahian, F., & Dautenhahn, K. (2015). Would you trust a (faulty) robot? *Proceedings of the Tenth Annual ACM/IEEE International Conference on Human-Robot Interaction*, 141–148. <https://doi.org/10.1145/2696454.2696497>
- Sarkar, S., Araiza-Illan, D., & Eder, K. (2017). Effects of faults, experience, and personality on trust in a robot co-worker. *ArXiv Preprint*, 1703, 02335. <http://arxiv.org/abs/1703.02335>
- Schaefer, K. E., Chen, J. Y. C., Szalma, J. L., & Hancock, P. A. (2016). A meta-analysis of factors influencing the development of trust in automation: Implications for understanding autonomy in future systems. *Human Factors*, 58(3), 377–400. <https://doi.org/10.1177/0018720816634228>
- Schmidt, P., & Biessmann, F. (2020). Calibrating human-AI collaboration: impact of risk, ambiguity and transparency on algorithmic bias. In A. H. et al. (Eds.), *Machine learning and knowledge extraction. CD-MAKE 2020. Lecture notes in computer science* (Vol. 12279, pp. 431–449). Springer. https://doi.org/10.1007/978-3-030-57321-8_24
- 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, 1–10. <https://doi.org/10.1016/j.ijhcs.2020.102551>
- Schaefer, K. E., & Scribner, D. R. (2015). Individual differences, trust, and vehicle autonomy. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 59(1), 786–790. <https://doi.org/10.1177/1541931215591242>
- Schaefer, K. E., & Straub, E. R. (2016). Will passengers trust driverless vehicles? Removing the steering wheel and pedals. *IEEE International Multi-Disciplinary Conference on Cognitive Methods in Situation Awareness and Decision Support (CogSIMA)*, 2016, 159–165. <https://doi.org/10.1109/COGSIMA.2016.7497804>
- Schrum, M. L., Neville, G., Johnson, M., Moorman, N., Paleja, R., Feigh, K. M., & Gombolay, M. C. (2021). Effects of social factors and team dynamics on adoption of collaborative robot autonomy. *ACM/IEEE International Conference on Human-Robot Interaction*, 149–157. <https://doi.org/10.1145/3434073.3444649>
- Siau, K., & Wang, W. (2018). Building trust in artificial intelligence, machine learning, and robotics. *Cutter Business Technology Journal*, 31(2), 47–53.
- Sindermann, C., Riedl, R., & Montag, C. (2020). Investigating the relationship between personality and technology acceptance with a focus on the smartphone from a gender perspective: Results of an exploratory survey study. *Future Internet*, 12(7), 110. <https://doi.org/10.3390/fi1207011>
- Söllner, M., Hoffmann, A., Hoffmann, H., Wacker, A., & Leimeister, J. M. (2012). Understanding the formation of trust in IT artifacts. *ICIS 2012 Proceedings*. 11. <https://aisel.aisnet.org/icis2012/proceedings/HumanBehavior/11>
- Sorrentino, A., Mancoppi, G., Coviello, L., Cavallo, F., & Fiorini, L. (2021). Feasibility study on the role of personality, emotion, and engagement in socially assistive robotics: A cognitive assessment scenario. *Informatics*, 8(2), 23. <https://doi.org/10.3390/informatics8020023>
- Sutter, M., & Kocher, M. G. (2007). Trust and trustworthiness across different age groups. *Games and Economic Behavior*, 59(2), 364–382. <https://doi.org/10.1016/j.geb.2006.07.006>
- Szalma, J. L., & Taylor, G. S. (2011). Individual differences in response to automation: The five factor model of personality. *Journal of Experimental Psychology: Applied*, 17(2), 71–96. <https://doi.org/10.1037/a0024170>
- Tapus, A., Țăpuș, C., & Matarić, M. J. (2008). User—Robot personality matching and assistive robot behavior adaptation for post-stroke rehabilitation therapy. *Intelligent Service Robotics*, 1, 169–183. <https://doi.org/10.1007/s11370-008-0017-4>
- Tenhundfeld, N. L., de Visser, E. J., Ries, A. J., Finomore, V. S., & Tossell, C. C. (2020). Trust and distrust of automated parking in a tesla model X. *Human Factors*, 62(2), 194–210. <https://doi.org/10.1177/0018720819865412>
- Thiebes, S., Lins, S., & Sunyaev, A. (2021). Trustworthy artificial intelligence. *Electronic Markets*, 31(2), 447–464. <https://doi.org/10.1007/s12525-020-00441-4>
- Tong, S. T., Corriero, E. F., Matheny, R. G., & Hancock, J. T. (2018). Online daters' willingness to use recommender technology for mate selection decisions. *IntRS Workshop*, 2225, 45–52.
- Tupes, E. C., & Christal, R. E. (1961). Recurrent personality factors based on trait ratings. *USAF ASD tech. Rep.*, 60(61–97), 225–251. <https://doi.org/10.1111/j.1467-6494.1992.tb00973.x>
- vom Brocke, J., Simons, A., Niehaves, B., Reimer, K., Plattfaut, R., & Cleven, A. (2009). Reconstructing the giant: On the importance of rigour in documenting the literature search process. *Proceedings of the European Conference on Information Systems*, 2206–2217.
- vom Brocke, J., Hevner, A., Léger, P. M., Walla, P., & Riedl, R. (2020). Advancing a NeuroIS research agenda with four areas of societal contributions. *European Journal of Information Systems*, 29(1), 9–24. <https://doi.org/10.1080/0960085X.2019.1708218>
- Voinescu, A., Morgan, P. L., Alford, C., & Caleb-Solly, P. (2018). Investigating older adults' preferences for functions within a human-machine interface designed for fully autonomous vehicles. In Zhou, J., & Salvendy, G. (Ed.), *Human aspects of IT for the aged population. applications in health, assistance, and entertainment. ITAP 2018. Lecture notes in computer science* (Vol. 10927, pp. 445–462). Springer. https://doi.org/10.1007/978-3-319-92037-5_32
- Walker, F., Wang, J., Martens, M. H., & Verwey, W. B. (2019). Gaze behaviour and electrodermal activity: Objective measures of drivers' trust in automated vehicles. *Transportation Research Part F: Traffic Psychology and Behaviour*, 64, 401–412. <https://doi.org/10.1016/j.trf.2019.05.021>
- Walsham, G. (1995). The emergence of interpretivism in IS research. *Information Systems Research*, 6(4), 376–394. <https://doi.org/10.1287/isre.6.4.376>
- Webster, J., & Watson, R. T. (2002). Analyzing the past to prepare for the future: Writing a literature review. *MIS Quarterly*, 26(2), xiii–xxiii.
- Xu, K. (2019). First encounter with robot Alpha: How individual differences interact with vocal and kinetic cues in users' social responses. *New Media & Society*, 21(11–12), 2522–2547. <https://doi.org/10.1177/1461444819851479>
- Yang, F., Huang, Z., Scholtz, J., & Arendt, D. L. (2020). How do visual explanations foster end users' appropriate trust in machine learning? In *25th International Conference on Intelligent User Interfaces (IUI '20)* (pp. 189–201). <https://doi.org/10.1145/3377325.3377480>
- Yorita, A., Egerton, S., Oakman, J., Chan, C., & Kubota, N. (2019). Self-adapting chatbot personalities for better peer support. *IEEE international conference on systems, man and cybernetics* (pp. 4094–4100) <https://ieeexplore.ieee.org/document/8914583>
- Youn, S., & Jin, S. V. (2021). “In A.I. we trust?” The effects of parasocial interaction and technopian versus luddite ideological views on chatbot-based customer relationship management in the emerging “feeling economy”. *Computers in Human Behavior*, 119, 106721. <https://doi.org/10.1016/j.chb.2021.106721>
- Zalake, M. (2020). Advisor: Agent-based intervention leveraging individual differences to support mental wellbeing of college students. *CHI EA '20: Extended abstracts of the 2020 CHI conference on human factors in computing systems*, 1–8. <https://doi.org/10.1145/3334480.3375026>
- Zhang, P. (2013). The affective response model: A theoretical framework of affective concepts and their relationships in the ICT context. *MIS Quarterly*, 37(1), 247–274 <https://www.jstor.org/stable/43825945>
- Zhang, T., Tao, D., Qu, X., Zhang, X., Zeng, J., Zhu, H., & Zhu, H. (2020). Automated vehicle acceptance in China: Social influence and initial trust are key determinants. *Transportation Research Part C: Emerging Technologies*, 112, 220–233. <https://doi.org/10.1016/j.trc.2020.01.027>
- Zhou, J., Luo, S., & Chen, F. (2020). Effects of personality traits on user trust in human-machine collaborations. *Journal on Multimodal User Interfaces*, 14(4), 387–400. <https://doi.org/10.1007/s12193-020-00329-9>
- Zimbardo, P., Johnson, R., & McCann, V. (2021). *Psychology: Core concepts (8th edition)*. Pearson.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.