




Systematic Review

Decoding Trust in Artificial Intelligence: A Systematic Review of Quantitative Measures and Related Variables

Letizia Aquilino ^{1,*}, Cinzia Di Dio ^{1,2}, Federico Manzi ^{1,2,3,4}, Davide Massaro ^{1,2}, Piercosma Bisconti ^{5,6} and Antonella Marchetti ^{1,2}

¹ Research Centre of Theory of Mind and Social Competences, Department of Psychology, Università Cattolica del Sacro Cuore, 20123 Milan, Italy; cinzia.didio@unicatt.it (C.D.D.); federico.manzi@unicatt.it (F.M.); davide.massaro@unicatt.it (D.M.); antonella.marchetti@unicatt.it (A.M.)

² Research Unit in Psychology and Robotics in the Life-Span, Department of Psychology, Università Cattolica del Sacro Cuore, 20123 Milan, Italy

³ IRCCS Fondazione Don Carlo Gnocchi, 20148 Milan, Italy

⁴ Intelligent Robotics Lab., Department of System Innovation, Osaka University, Osaka 560-8531, Japan

⁵ DEXAI Etica Artificiale, 00179 Rome, Italy; piercosma.bisconti@dexai.eu or bisconti@diag.uniroma1.it

⁶ DIAG, Department of Computer, Control, and Management Engineering, Sapienza University, 00185 Rome, Italy

* Correspondence: letizia.aquilino1@unicatt.it

Abstract

As artificial intelligence (AI) becomes ubiquitous across various fields, understanding people's acceptance and trust in AI systems becomes essential. This review aims to identify quantitative measures used to measure trust in AI and the associated studied elements. Following the PRISMA guidelines, three databases were consulted, selecting articles published before December 2023. Ultimately, 45 articles out of 1283 were selected. Articles were included if they were peer-reviewed journal publications in English reporting empirical studies measuring trust in AI systems with multi-item questionnaires. Studies were analyzed through the lenses of cognitive and affective trust. We investigated trust definitions, questionnaires employed, types of AI systems, and trust-related constructs. Results reveal diverse trust conceptualizations and measurements. In addition, the studies covered a wide range of AI system types, including virtual assistants, content detection tools, chatbots, medical AI, robots, and educational AI. Overall, the studies show compatibility of cognitive or affective trust focus between theorization, items, experimental stimuli, and level of anthropomorphism of the systems. The review underlines the need to adapt measurement of trust in the specific characteristics of human–AI interaction, accounting for both the cognitive and affective sides. Trust definitions and measurement could be chosen depending also on the level of anthropomorphism of the systems and the context of application.

Keywords: artificial intelligence; perceived trustworthiness; trust; machine learning



Academic Editor: Roberto Theron

Received: 29 May 2025

Revised: 3 July 2025

Accepted: 7 July 2025

Published: 14 July 2025

Citation: Aquilino, L.; Di Dio, C.; Manzi, F.; Massaro, D.; Bisconti, P.; Marchetti, A. Decoding Trust in Artificial Intelligence: A Systematic Review of Quantitative Measures and Related Variables. *Informatics* **2025**, *12*, 70. <https://doi.org/10.3390/informatics12030070>

Copyright: © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Artificial intelligence is extensively used in an increasing number of areas and contexts in everyday life. Additionally, users are now engaging with tools that clearly disclose the use of AI. For this reason, it is important to explore the conditions under which these types of systems are accepted. Among these factors, trust in AI plays a particularly prominent role [1–3], as it has been shown to influence continued system adoption and perceptions of their usefulness and ease of use [4,5]. Trust is also a complex construct, with prior literature reviews revealing significant variability in the ways it has been conceptualized [6].

This complexity is further compounded by the fact that AI systems vary greatly in their modalities and areas of application, for example, considering the distinct features that differentiate an online shopping assistant from diagnostic systems in the medical field. Presumably due to this variety, there is also substantial variability in how trust is measured and in the factors studied in relation to it [4,5].

One possible way to further the understanding of trust processes is to adopt a conceptualization that enables us to grasp its complexity. To that end, this review will employ the theory that defines trust as comprising two components: cognitive and affective. Trust, in fact, develops from the combined action of rational, deliberate thoughts and feelings, intuition, and instinct [7].

Cognitive trust represents a rational, evidence-based route of trust formation, centring on an individual's evaluation of another's reliability and competence. This trust is grounded in an assessment of the trustee's knowledge, skills, or predictability, aligning with the notion of rational judgement in interpersonal relationships [8–10]. These attributes of being reliable and predictable represent the trustee's trustworthiness, which derives from their evaluation by the trustor [11]. Several scholars categorize credibility or reliability as cognitive trust, contrasting it with affective trust, which is rooted in emotional bonds [9,12–14]. Cognitive trust involves an analysis-based approach where individuals use evidence and logical reasoning to assess the trustee's dependability, differentiating between trustworthy, distrusted, and unknown parties [7]. Within this process, cognitive trust mirrors the attitude construct, as it is grounded in evaluative processes that rely on "good reasons" to trust [7]. In human relationships, cognitive trust reflects a person's confidence in the trustee's competence. This evolves according to the knowledge accumulated in the trustor–trustee relationship, which helps to form predictions about the likelihood of their consistency and adherence to commitments [10,15]. The development of cognitive trust is therefore gradual, requiring a foundation of knowledge and familiarity that allows each side to estimate trustworthiness based on the situational evidence available within the relationship [16,17].

Nevertheless, it is important to emphasize that trust in someone is not just a matter of rational estimation of their trustworthiness, but is closely linked to affective trust involving aspects such as mutual care and emotional bonding that, if good enough, can generate feelings of security within a relationship [10,18,19]. Unlike cognitive trust, which is grounded in rational evaluation, affective trust is shaped by personal experiences and the emotional investment partners make, often resulting in trust that extends beyond what is justified by concrete knowledge [13,20]. This type of trust relies on the perception of a partner's genuine care and intrinsic motivation, forming a confidence based on emotional connections [10]. Affective trust is particularly relevant in settings where kindness, empathy, and bonding are perceived as core aspects of a relationship [14,21]. Within affective trust, emotional security provides assurance and comfort, helping individuals believe that expected benefits will materialize even without actual validation [22]. This form of trust plays a significant role in relationships, often prevailing as a strong indicator of overall perceived trust [23,24]. The affective response embodies these "emotional bonds" [17], complementing cognitive trust with an instinctive feeling of faith and security, even amid future uncertainties [7,10].

Although general trust theories were first developed to explain human relationships, their application to human–AI interaction can be justified in light of the peculiar characteristics these systems hold. AI technologies have, in fact, features that allow them to respond in a syntonetic and consistent manner with user requests. This can be perceived as more human-like than others and thus elicit modes of interactions that presuppose social rules and norms typical of humans relationships [25–28]. Nevertheless, not all systems present the same characteristics. As some can be equipped with clear human-like features, such

as virtual assistants that use natural language to communicate, others do not resemble humans at all, for example, AI-powered machines operating in factories.

We argue that, based on the specific characteristics of each AI system, the two trust components may have different degrees of relevance. For example, the constant use and presence of features that encourage the anthropomorphism of the system could lead to trust being more strongly explained by its affective side. In fact, anthropomorphism involves attributing human characteristics to artificial agents [29,30], which, along with regular interaction, is hypothesized to foster an emotional connection. This could cause the user to rely on the technology not so much because of a careful and rational assessment of its ability to achieve a goal, but rather due to the parasocial relationship created with it over time.

Another factor that could be considered to hypothesize the dominance of one component of trust over another is the level of risk associated with choosing to rely on the technological tool. An example of a low-risk application can be found in Netflix's recommendation algorithm, which helps viewers to decide what to watch based on their preferences [31]. In other cases, safety and survival of the person relying on the system are involved, as it is with self-driving cars or medical AI [32,33]. It is possible to hypothesize that the cognitive component of trust, more than the affective one, is primarily involved in high-risk applications than the affective one, where a person's self-preservation instinct requires that the choice be marked by a good degree of certainty regarding the likelihood of success.

In light of this complexity associated with the trust construct in human relationships, it becomes even more important to explore the cognitive and affective components in the relationship with AI. This paper aims to identify how these components have been analyzed within research on trust in AI. In particular, it aims to systematize the methodological elements by highlighting critical aspects of self-report measures widely used for trust analysis.

The novelty of the current work consists in providing parameters that could guide future research in a more coherent and comparable way. Previous reviews [6,34] have shown great methodological variability, which often entails replicability issues. These could be overcome by employing specific criteria to decide on how to measure trust in artificial systems, specifically through the lens of affective–cognitive trust theory. Accordingly, it will consider definitions or conceptualizations of the trust construct (cognitive and affective), the characteristics of the scales used to measure it, the experimental stimuli, the types of systems studied, and the variables examined in association with trust.

2. Materials and Methods

A systematic review of the scientific literature was conducted to examine how trust in AI is measured in its cognitive and affective components using self-report scales. Only journal articles employing self-report scales with more than one item were considered eligible for inclusion. A review protocol was compiled, following the Preferred Reporting Items for Systematic Reviews and Meta-Analysis (PRISMA) guidelines [35].

2.1. Study Selection Criteria

The following selection criteria were applied to the articles found in databases: research studies were eligible for inclusion; chapters, books, reviews, conference papers, and workshop papers were excluded. Acknowledging the relevance of studies presented and published in conference and workshop proceedings in technical fields, this choice was applied to guarantee a satisfactory psychometric methodological quality. An argument is made that even among research conducted and analyzed in other fields of study, such as computer science, engineering or economics, undergoing a tied peer-review process, as the

one required by journals, could be indication of more rigorous human-related measures. The structural requirements for conference and workshop papers also usually entail brevity, which in turn allows them to contain extensive methodological information (e.g., [36,37]). They would therefore be unsuited for the current work as they do not provide the necessary material to analyze.

The abstracts of the identified publications were screened for relevance to the selection criteria. Specific inclusion criteria were measuring trust in AI systems and using self-report questionnaires consisting of more than one item to measure trust. To that end, papers that described qualitative or mixed methods or quantitative methods with an insufficient number of items were rejected. Although no restrictions were applied with regard to the scientific field of study, the aim was to gather measures that could be psychometrically valid. To this end, the multi-item criterion was introduced, as it has been demonstrated how this kind of scale can guarantee a higher level of both reliability (by adjusting random error and increasing accuracy [38,39]) and construct validity, especially in the case of multifaceted constructs [40,41] such as trust. We included only journal articles in English. There were no restrictions regarding the type of AI system under scope and the field of application of the technology. This review examines scientific literature published before December 2023. It can be seen, however, that almost all the articles collected by the research team were produced in the last five years. This is consistent with the trend in the development of AI system functions. Modern AI interfaces feature greater interactivity and usability characteristics [42,43], which have contributed to their increasing diffusion in recent years. As a consequence, the interest of research on the topic has also increased, as has the specific focus on aspects such as fairness [44,45]. This temporal distribution ensures the analysis of articles that address the most recent and currently utilized forms of systems present in users' daily lives.

2.2. Data Sources and Search Strategy

Electronic literature searches were performed using Web of Science, ACM Digital Library, and Scopus. Two researchers reviewed the potential studies individually for eligibility. A list of keywords was used to identify the studies, through an interactive process of search and refine. Table 1 shows the detailed search strategy. Each database was searched independently, according to a specific interaction research string: ("Trust" OR "Trustworthiness") AND ("AI" OR "Artificial Intelligence" OR "Machine Learning"). A choice was made to include both "trust" and "trustworthiness" given the close relationship between the two, with the latter being one of the main antecedents of the first, and with the aim of exploring measurement that encompasses multiple phases of the trusting process, from perception of trustworthiness to the actual trust and consequent trusting intention. No limitation regarding research areas was applied. The article screening process was conducted using the Rayyan platform.

Table 1. Detailed search strategy accounting for the subtotal of publications for each database; the quantity of articles removed for specific reasons (duplicates between different databases and wrong type of publication, namely not journal articles); the number of articles screened for title and abstract; and finally the number of articles included in the analysis.

Trust OR Trustworthiness AND	Web of Science	Scopus	ACM Digital Library
Artificial Intelligence	162	194	28
AI	222	298	52
Machine Learning	133	135	59
Subtotal	517	627	139

Table 1. Cont.

Trust OR Trustworthiness AND	Web of Science	Scopus	ACM Digital Library
Total			1283
Duplicated removal			360
Wrong type of publication			281
Identified studies for Abstract and Title screening			642
Included			45

3. Results

After the removal of duplicates and title and abstract screening of electronic database search results, 50 articles were full-text-screened. Five articles showed a methodology that did not meet the criteria. For this reason, a total of 45 articles were included (see Table 2 for the summary of the studies). Figure 1 shows the study selection flow chart. The results are organized with respect to the following sections: definitions of trust employed, types of AI systems considered, experimental stimulus, and most significant constructs related to trust in AI.

Table 2. Papers included: brief description.

N	Authors	Research Questions (RQ)/Hypotheses (H) Regarding Trust	Results
1	[46]	RQ1: What is the difference in the level of trust in AI tools for spam review detection between younger and older adults? RQ2: How does the difference in credibility judgments of reviews between humans and AI tools affect the users' trust in AI tools?	Older adults' evaluations of the competence, benevolence, and integrity of AI tools were found to be far higher than younger adults'.
2	[47]	RQ1: Does providing information about the performance of AI enhance or harm users' levels of (a) cognitive and (b) behavioural trust in AI?	Regardless of information framing, higher levels of trust were perceived by those who received no information about the AI's performance; participants were more likely to perceive high levels of trust when they did not feel ownership of the message.
3	[48]	RQ1: Does AI visibility result in lower trust during the process of purchasing health insurance online? RQ2: Does perceived ease of use have the same influence on trust in AI if AI is visible during the purchase of health insurance online?	Trust was higher without visible AI involvement; perceived ease of use had the same influence on trust in AI, regardless of AI visibility.
4	[49]	RQ1: What are the anthropomorphic attributes of chatbots influencing consumers' perceived trust in and responses to chatbots? RQ2: How do the anthropomorphic attributes of chatbots influence consumers' perceived trust in and subsequent responses to chatbots? RQ3: To what extent do the impacts of the anthropomorphic attributes of chatbots on consumers' perceived trust in chatbots depend on the type of relationship norm that is salient in consumers' minds during the service encounter?	Chatbot anthropomorphism has a significant effect on consumers' trust in chatbots (when consumers infer chatbots as having higher warmth or higher competence based on interaction with them, they tend to develop higher levels of trust in them). Among the three anthropomorphic attributes, perceived competence showed the largest effect size. The effect of the anthropomorphic attributes of chatbots on consumers' trust in them resulted to be contingent on the type of relationship norm that is salient in their minds during the service encounter.

Table 2. Cont.

N	Authors	Research Questions (RQ)/Hypotheses (H) Regarding Trust	Results
5	[50]	How do citizens evaluate the impact of AI decision-making in government in terms of impact on citizen values (trust)?	Complexity was not shown to have a relevant effect on trust.
6	[51]	RQ1: Compared with no transparency, what are the effects of placebic transparency on (a) uncertainty, (b) trust in judgments, (c) trust in system, and (d) use intention?	Machine-agency locus induced less social presence, which in turn was associated with lower uncertainty and higher trust. Transparency reduced uncertainty and enhanced trust.
7	[52]	RQ1: What is the relationship between a user's trust in a VAS and intention to use the VAS? RQ2. What is the relationship between a VAS's interaction quality and a user's trust in the VAS?	Interaction quality significantly led to user trust and intention to use
8	[53]	RQ1: What is the impact of social influence on perceived trust? RQ2: What is the impact of perceived trust on the intention to use AI? RQ3: What is the mediation impact of personal innovativeness and perceived trust between social influence and the intention to use AI?	Social influence had a significant direct positive effect on perceived trust, and perceived trust had a significant direct positive effect on the intention to use. Personal innovativeness and perceived trust partially mediated social influence and the intention to use.
9	[4]	RQ1: How do explainability and causability affect trust and the user experience with a personalized recommender system?	Causability played an antecedent role to explainability and an underlying part in trust. Users acquired a sense of trust in algorithms when they were assured of their expected level of FATE. Trust significantly mediated the effects of the algorithms' FATE on users' satisfaction. Satisfaction stimulated trust and in turn led to positive user perception of FATE. Higher satisfaction led to greater trust and suggested that users were more likely to continue to use an algorithm.
10	[54]	RQ1: What is the effect of AT ethicality on trust within human–AI teams? RQ2: If unethical actions damage trust, how effective are common trust repair strategies after an AI teammate makes an unethical decision?	AT's unethical actions decreased trust in the AT and the overall team. Decrease in trust in an AT was not associated with decreased trust in the human teammate.
11	[55]	RQ1: What influence, if any, do the ethical requirements of AI have on trust in AI?	Younger and more educated individuals indicated greater trust in AI, as did familiarity with smart consumer technologies. Propensity for trust in other people as well as trust in institutions were strongly correlated with both dimensions of trust (human-like and functionality trust).
12	[1]	RQ1: Is there a difference in the influence between the human-like dimension and the functionality dimension of trust in AI within the TAM framework?	Trust had a significant effect the on intention to use AI. Both dimensions of trust (human-like and functionality) shared a similar pattern of effects within the model, with functionality-related trust exhibiting a greater total impact on usage intention than human-like trust

Table 2. Cont.

N	Authors	Research Questions (RQ)/Hypotheses (H) Regarding Trust	Results
13	[56]	RQ1: When humans and AI together serve as moderators of content (vs. AI only or human only), what is the relationship between interactive transparency (vs. transparency vs. no transparency) and (a) trust, (b) agreement, (c) understanding of the system and (d) perceived user agency, human agency, and AI agency?	Users trusted AI for the moderation of content just as much as humans, but it depended on the heuristic that was triggered when they were told AI was the source of moderation. Allowing users to provide feedback to the algorithm enhanced trust by increasing user agency.
14	[57]	RQ1: How does anthropomorphism affect customer trust in AI? RQ2: What effect does empathy response have on customer trust in AI? RQ3: What effect does interaction response have on customer trust in AI? RQ4: What is the relationship between communication quality and customer trust in AI?	Anthropomorphism and interaction did not play critical roles in generating customer trust in AI unless they created communication quality with customers.
15	[58]	RQ1: What is the influence of perceived usefulness of voice-activated assistants on users' attitude to use and trust towards the technology? RQ2: What is the influence of perceived ease of use of voice-activated assistants on users' attitude to use and trust towards the technology? RQ3: What is the influence of perceived enjoyment of voice-activated assistants on users' attitude to use and trust towards the technology? RQ4: What is the influence of perceived social presence of voice-activated assistants on users' attitude to use and trust towards the technology? RQ5: What is the influence of user-inferred social cognition of voice-activated assistants on users' attitude to use and trust towards the technology? RQ6: What is the influence of perceived privacy concerns of voice-activated assistants on users' attitude to use and trust towards the technology?	The social attributes (social presence and social cognition) were the unique antecedents for developing trust. Additionally, a peculiar dynamic between privacy and trust was shown, highlighting how users distinguish two different sources of trustworthiness in their interactions with VAs, identifying the brand producers as the data collector.
16	[59]	RQ1: What is the effect of uncanniness on trust in artificial agents?	Uncanniness had a negative impact on trust in artificial agents. Perceived harm and perceived injustice were the major predictors of uncanniness.
17	[60]	RQ1: Does employees' perceived transparency mediate the impact of AI decision-making transparency on employees' trust in AI? RQ2: What impact does employees' perceived effectiveness of AI have on employees' trust in AI? RQ3: Do employees' perceived transparency and perceived effectiveness of AI have a chain mediating role between AI decision-making transparency and employees' trust in AI? RQ4: Does employees' discomfort with AI have a negative impact on employees' trust in AI? RQ5: Do employees' perceived transparency of and discomfort toward AI have a chain mediating role between AI decision-making transparency and employees' trust in AI?	AI decision-making transparency (vs. non-transparency) led to higher perceived transparency, which in turn increased both effectiveness (which promoted trust) and discomfort (which inhibited trust).

Table 2. Cont.

N	Authors	Research Questions (RQ)/Hypotheses (H) Regarding Trust	Results
18	[61]	<p>RQ1: How is AI trusted compared to human doctors?</p> <p>RQ2: When an AI system learns and then provides patients' desired treatment, will they perceive the AI system to care for them and, therefore, will they trust the AI system more than when the AI system provides patients' desired treatment without learning it?</p> <p>RQ3: Will an AI system proposing the treatment preferred by the patient be trusted more than that not proposing the patient's preferred treatment, regardless of whether or not it learns the patient's preference?</p>	<p>Participants trusted the AI system less than a doctor, even when the AI system learned and suggested their desired treatment and even if it performed at the level of a human doctor.</p>
19	[62]	<p>RQ1: How does trust in Siri influence brand loyalty?</p> <p>RQ2: How do interactions with Siri influence brand loyalty?</p> <p>RQ3: Will a higher level of perceived risk be associated with a lower level of brand loyalty?</p> <p>RQ4: How does the novelty value of Siri influence brand loyalty?</p>	<p>Perceived risk seemed to have a significantly negative influence on brand loyalty. The influence of novelty value of using Siri was found to be moderated by brand involvement and consumer innovativeness in such a way the influence is greater for consumers who are less involved with the brand and who are more innovative.</p>
20	[63]	<p>RQ1: Does employees' general trust in technology impact on their trust in AI in the company?</p> <p>RQ2: Does intra-organizational trust impact on employees' trust in AI in the company?</p> <p>RQ3: Does employees' individual competence trust impact their trust in AI in the company?</p>	<p>A positive relationship between general trust in technology and employees' trust in AI in the company was found, as well as between intraorganizational trust and employees' trust in AI in the company</p>
21	[64]	<p>RQ1: Does human-computer trust play a mediating role in the relationship between patients' self-responsibility attribution and acceptance of medical AI for independent diagnosis and treatment?</p> <p>RQ2: Does human-computer trust play a mediating role in the relationship between patients' self-responsibility attribution and acceptance of medical AI for assistive diagnosis and treatment?</p> <p>RQ3: Do Big Five personality traits moderate the relationship between human-computer trust and acceptance of medical AI for independent diagnosis and treatment?</p>	<p>Patients' self-responsibility attribution was positively related to human-computer trust (HCT). Conscientiousness and openness strengthened the association between HCT and acceptance of AI for independent diagnosis and treatment; agreeableness and conscientiousness weakened the association between HCT and acceptance of AI for assistive diagnosis and treatment.</p>

Table 2. Cont.

N	Authors	Research Questions (RQ)/Hypotheses (H) Regarding Trust	Results
22	[65]	<p>RQ1: Does human–computer trust mediate the relationship between performance expectancy and healthcare workers’ adoption intention of AI-assisted diagnosis and treatment?</p> <p>RQ2: Does human–computer trust mediate the relationship between effort expectancy and healthcare workers’ adoption intention of AI-assisted diagnosis and treatment?</p> <p>RQ3: Can the relationship between performance expectancy and healthcare workers’ adoption intention of AI-assisted diagnosis and treatment be mediated sequentially by social influence and human–computer trust?</p> <p>RQ4: Can the relationship between effort expectancy and healthcare workers’ adoption intention of AI-assisted diagnosis and treatment be mediated sequentially by social influence and human–computer trust?</p>	<p>Social influence and human–computer trust, respectively, mediated the relationship between expectancy (performance expectancy and effort expectancy) and healthcare workers’ adoption intention of AI-assisted diagnosis and treatment. Furthermore, social influence and human–computer trust played a chain mediation role between expectancy and healthcare workers’ adoption intention of AI-assisted diagnosis and treatment.</p>
23	[66]	<p>RQ1: Will an interactive process (Interactive Machine Learning) develop behaviours that are more trustworthy and adhere more closely to user goals and expectations about search performance during implementation?</p>	<p>Compared to noninteractive techniques, Interactive Machine Learning (IML) behaviours were more trusted and preferred, as well as recognizable, separate from non-IML behaviours.</p>
24	[67]	<p>RQ1: What is the influence of trust on clinicians based on the ML explanations?</p>	<p>While the clinicians’ trust in automated diagnosis increased with the explanations, their reliance on the diagnosis reduced as clinicians were less likely to rely on algorithms that were not close to human judgement.</p>
25	[68]	<p>RQ1: Do explainability artifacts increase user confidence in a black-box AI system?</p>	<p>Users’ trust of black-box systems was high and explainability artifacts did not influence this behaviour</p>
26	[69]	<p>RQ1: How do teachers’ perceived trust affect their acceptance of EAITs?</p> <p>RQ2: What is the most dominant factor that affects teachers’ intention to use EAITs?</p>	<p>Teachers with constructivist beliefs were more likely to integrate EAITs than teachers with transmissive orientations. Perceived usefulness, perceived ease of use, and perceived trust in EAITs were determinants to be considered when explaining teachers’ acceptance of EAITs. The most influential determinant of predicting their acceptance was found to be how easily the EAIT is constructed.</p>

Table 2. Cont.

N	Authors	Research Questions (RQ)/Hypotheses (H) Regarding Trust	Results
27	[70]	<p>RQ1: What is the correlation between users' behaviour of trusting AI virtual assistants and the acceptance of AI virtual assistants?</p> <p>RQ2: What is the correlation between perceived usefulness and users' behaviour of trusting AI virtual assistants?</p> <p>RQ3: What is the correlation between perceived ease of use and users' behaviour of trusting AI virtual assistants?</p> <p>RQ4: What is the relationship between perceived humanity and user trust in artificial intelligence virtual assistants?</p> <p>RQ5: What is the correlation between perceived social interactivity and user trust in artificial intelligence virtual assistants?</p> <p>RQ6: What is the correlation between perceived social presence and user trust in AI virtual assistants?</p> <p>User trust behaviour has a mediating role between functionality and acceptance.</p> <p>H5: User trust behaviour has a mediating role between social emotion and acceptance.</p>	<p>Functionality and social emotions had a significant effect on trust, where perceived humanity showed an inverted U-shaped relationship with trust, and trust mediated the relationship between both functionality and social emotions and acceptance.</p>
28	[71]	<p>RQ1. How do different explanations affect user's trust in AI systems?</p> <p>RQ2. How do the levels of risk of the user's decision-making play a role in the user's trust in AI systems?</p> <p>RQ3. How does the performance of the AI system play a role in the user's AI trust even when an explanation is present?</p>	<p>The study has shown that the presence of explanations increases AI trust, but only in certain conditions. AI trust was higher when explanations with feature importance were provided than with counterfactual explanations. Moreover, when the system performance is not guaranteed, the use of explanations seems to lead to an overreliance on the system. Lastly, system performance had a stronger impact on trust, compared to the effects of other factors (explanation and risk).</p>
29	[72]	<p>H1. Trust in AI has a positive effect on the degree of use of various AI-based services.</p>	<p>First, trust in AI and privacy-protective behaviour positively impact AI-based service usage. Second, online skills do not impact AI-based service usage significantly.</p>
30	[73]	<p>H1: The different anthropomorphism of social robots affects students' initial capability trust in social robots.</p> <p>H3: The stronger the student's attraction perception of the social robot, the stronger the trust in the initial capabilities of the social robot.</p> <p>H5: There are differences in the initial capability of students of different ages to trust the initial capabilities of social robots with different anthropomorphic levels.</p>	<p>When the degree of anthropomorphism of social robots is at different levels, there are significant differences in students' initial capability trust evaluation. It can be seen that the degree of anthropomorphism of social robots has an impact on students' initial capability trust.</p>

Table 2. Cont.

N	Authors	Research Questions (RQ)/Hypotheses (H) Regarding Trust	Results
31	[74]	RQ3. Does trust in AI play a role informing consumers' intention to purchase using AI?	The results show that consumers tend to demand anthropomorphized products to gain a better shopping experience and, therefore, demand features that attract and motivate them to purchase through artificial intelligence via mediating variables, such as perceived animacy and perceived intelligence. Moreover, trust in artificial intelligence moderates the relationship between perceived anthropomorphism and perceived animacy.
32	[75]	H2-1: Information seeking positively affects the perceived trust of generative AI. H2-2: Task efficiency positively affects the perceived trust of generative AI. H2-3: Personalization positively affects the perceived trust of generative AI. H2-4: Social interaction positively affects the perceived trust of generative AI. H2-5: Playfulness positively affects the perceived trust of generative AI. H4: Perceived trust of generative AI positively affects continuance intention.	The findings reveal a negative relationship between personalization and creepiness, while task efficiency and social interaction are positively associated with creepiness. Increased levels of creepiness, in turn, result in decreased continuance intention. Furthermore, task efficiency and personalization have a positive impact on trust, leading to increased continuance intention.
33	[76]	RQ1: Is there an initial difference for trustworthiness assessments, trust, and trust behaviour between the automated system and the human trustee? RQ2: Is there an initial difference and are there different effects for trust violations and trust repair interventions for the facets of trustworthiness for human and automated systems as trustees? RQ3: Will there be interaction effects between the trust repair intervention and the information regarding imperfection for the automated system for trustworthiness, trust, and trust behaviour?	The results of the study showed that participants have initially less trust in automated systems. Furthermore, the trust violation and the trust repair intervention had weaker effects for the automated system. Those effects were partly stronger when highlighting system imperfection.
34	[77]	H1: The higher the perceived reliability of AI is, the more cognitive trust employees display in AI. H2: The higher the AI transparency is, the more cognitive trust employees display in AI. H3: The higher AI flexibility is, the more cognitive trust employees have in AI. H4: Effectiveness of data governance stimulates trust in data governance, which leads to cognitive trust in AI. H5: AI-driven disruption in work routines lowers the effect of AI reliability, transparency, and flexibility on cognitive trust in AI.	The findings suggest that AI features positively influence the cognitive trust of employees, while work routine disruptions have a negative impact on cognitive trust in AI. The effectiveness of data governance was also found to facilitate employees' trust in data governance and, subsequently, employees' cognitive trust in AI.

Table 2. Cont.

N	Authors	Research Questions (RQ)/Hypotheses (H) Regarding Trust	Results
35	[78]	<p>H1a. Seeking information about AI on traditional media is positively associated with trust in AI after controlling for faith in and engagement with AI.</p> <p>H1b. Seeking information about AI on social media is negatively associated with trust in AI after controlling for faith in and engagement with AI.</p> <p>H2a. Concern about misinformation online will weaken the positive relationship between information-seeking behaviour about AI on traditional media and trust in AI.</p> <p>H2b. Concern about misinformation online will strengthen the negative relationship between information-seeking behaviour about AI on social media and trust in AI.</p>	<p>Results indicate a positive relationship exists between seeking AI information on social media and trust across all countries. However, for traditional media, this association was only present in Singapore. When considering misinformation, a positive moderation effect was found for social media in Singapore and India, whereas a negative effect was observed for traditional media in Singapore.</p>
36	[79]	<p>H1a: Accuracy experience of AI technology is positively related to the trust in AI-enabled services.</p> <p>H1b: Insight experience of AI technology is positively related to the trust in AI-enabled services.</p> <p>H1c: Interactive experience of AI technology is positively related to the trust in AI-enabled services.</p> <p>H3a: User trust is positively related to eWOM engagement.</p> <p>H3b: User trust is positively related to usage intention to use AI-enabled services.</p> <p>H5a: Self-efficacy positively moderates the association between accuracy experience and trust in AI.</p> <p>H5b: Self-efficacy positively moderates the association between insight experience and trust in AI.</p> <p>H5c: Self-efficacy positively moderates the association between interactive experience and trust in AI.</p> <p>H6: eWOM engagement mediates the association between trust in AI and AI-enabled service usage intention.</p>	<p>The results indicate that accuracy experience, insight experience, and interactive experience between trust in AI and eWOM engagements are significant, except for the relationship between interactive experience and eWOM. Likewise, the outcome indicates that trust in AI has a significant positive relationship between usage intention and eWOM, while eWOM significantly and positively influences usage intention. In addition, this study found that word of mouth mediates the association between accuracy experience and trust in AI. The results show that self-efficacy moderates the association between accuracy experience and trust in AI.</p>
37	[80]	<p>RQ: How do initial perceptions of a financial AI assistant differ between people who report they (rather) trust and people who report they (rather) do not trust the system?</p>	<p>Comparisons between high-trust and low-trust user groups revealed significant differences in both open-ended and closed-ended answers. While high-trust users characterized the AI assistant as more useful, competent, understandable, and human-like, low-trust users highlighted the system's uncanniness and potential dangers. Manipulating the AI assistant's agency had no influence on trust or intention to use.</p>

Table 2. Cont.

N	Authors	Research Questions (RQ)/Hypotheses (H) Regarding Trust	Results
38	[81]	<p>H1. Perceived anthropomorphism has a positive effect on trust in AI teammates.</p> <p>H2. Perceived rapport has a positive effect on trust in AI teammates.</p> <p>H3. Perceived enjoyment has a positive effect on trust in AI teammates.</p> <p>H4. Peer influence has a positive effect on trust in AI teammates.</p> <p>H5. Facilitating conditions has a positive effect on trust in AI teammates.</p> <p>H6. Self-efficacy has a positive effect on trust in AI teammates.</p> <p>H8. Trust in AI teammates has a positive effect on intention to cooperate with AI teammates.</p>	<p>The results show that perceived rapport, perceived enjoyment, peer influence, facilitating conditions, and self-efficacy positively affect trust in AI teammates. Moreover, self-efficacy and trust positively relate to the intention to cooperate with AI teammates.</p>
39	[82]	<p>H1: The expertise of AI chatbots has a positive impact on consumers' trust in chatbots.</p> <p>H2: The responsiveness of AI chatbots has a positive impact on consumers' trust in chatbots.</p> <p>H3: The ease of use of AI chatbots has a positive impact on consumers' trust in chatbots.</p> <p>H4: The anthropomorphism of AI chatbots has a positive impact on consumers' trust in chatbots.</p> <p>H5: Consumers' brand trust in AI chatbot providers positively affects consumers' trust in chatbots.</p> <p>H6: Human support has a negative impact on consumers' trust in AI chatbots.</p> <p>H7: Perceived risk has a negative impact on consumers' trust in AI chatbots.</p> <p>H8: Privacy concerns moderate the relationship between chatbot-related factors and consumers' trust in AI chatbots.</p> <p>H9: Privacy concerns moderate the relationship between company-related factors and consumers' trust in AI chatbots.</p>	<p>The results found that the chatbot-related factors (expertise, responsiveness, and anthropomorphism) positively affect consumers' trust in chatbots. The company-related factor (brand trust) positively affects consumers' trust in chatbots, and perceived risk negatively affect consumers' trust in chatbots. Privacy concerns have a moderating effect on company-related factors.</p>
40	[83]	<p>H1. Street-level bureaucrats perceive AI recommendations that are congruent with their professional judgement as more trustworthy than AI recommendations that are incongruent with their professional judgement.</p> <p>H2. Street-level bureaucrats perceive explained AI recommendations as more trustworthy than unexplained AI recommendations.</p> <p>H3. Higher perceived trustworthiness of AI recommendations by street-level bureaucrats is related to an increased likelihood of use.</p>	<p>We found that police officers trust and follow AI recommendations that are congruent with their intuitive professional judgement. We found no effect of explanations on trust in AI recommendations. We conclude that police officers do not blindly trust AI technologies, but follow AI recommendations that confirm what they already thought.</p>

Table 2. Cont.

N	Authors	Research Questions (RQ)/Hypotheses (H) Regarding Trust	Results
41	[84]	<p>H7. Trust positively affects performance expectancy of AI virtual assistants.</p> <p>H8. Trust positively affects behavioural intention to use AI virtual assistants.</p> <p>H9. Trust positively affects attitude toward using AI virtual assistants.</p> <p>H10. Trust negatively affects perceived risk of AI virtual assistants.</p>	<p>Results show that gender is significantly related to behavioural intention to use, education is positively related to trust and behavioural intention to use, and usage experience is positively related to attitude toward using. UTAUT variables, including performance expectancy, effort expectancy, social influence, and facilitating conditions, are positively related to behavioural intention to use AI virtual assistants. Trust and perceived risk respectively have positive and negative effects on attitude toward using and behavioural intention to use AI virtual assistants. Trust and perceived risk play equally important roles in explaining user acceptance of AI virtual assistants.</p>
42	[85]	<p>H2a. Consumers have greater trust in a high-intelligence-level AI compared to a low-intelligence-level AI.</p> <p>H2b. Trust mediates the relationship between the intelligence level of AI and consumers' willingness to delegate tasks to the AI.</p> <p>H3. Task objectivity has a moderating effect on the relationship between intelligence level of AI and trust. Specifically, consumers have greater trust in low-intelligence-level AI for objective tasks compared to subjective tasks, but such differences in trust are not manifested when the AI has a high intelligence level.</p> <p>H4. Anthropomorphism has a moderating effect on the relationship between intelligence level of AI and trust. Specifically, consumers' trust in AI may be enhanced by increased anthropomorphism when the AI's intelligence level is low, but this moderated effect is not manifested when its intelligence level is high.</p>	<p>The findings reveal that increasing AI intelligence enhances consumer decision delegation for both subjective and objective consumption tasks, with trust serving as a mediating factor. However, objective tasks are more likely to be delegated to AI. Moreover, anthropomorphic appearance positively influences trust but does not moderate the relationship between intelligence and trust.</p>
43	[86]	<p>H1b. Empathy response has positively impacted customer trust toward AI chatbots.</p> <p>H2a. Anonymity has positively impacted interaction.</p> <p>H2b. Anonymity has positively impacted customer trust toward AI chatbots.</p> <p>H5. Interaction has positively impacted customer trust toward AI chatbots.</p>	<p>The paper reports that empathy response, anonymity and customization significantly impact interaction. Empathy response is found to have the strongest influence on interaction. Meanwhile, empathy response and anonymity were revealed to indirectly affect customer trust.</p>

Table 2. Cont.

N	Authors	Research Questions (RQ)/Hypotheses (H) Regarding Trust	Results
44	[87]	<p>RQ1: How are various user characteristics, such as (a) demographics, (b) trust propensity, (c) fake news self-efficacy, (d) fact-checking service usage, (e) AI expertise, (f) and overall AI trust, associated with their trust in specific explainable ML fake news detectors?</p> <p>RQ2: How are the perceived machine characteristics such as (a) performance, (b) collaborative capacity, (c) agency, and (d) complexity associated with users' trust in explainable ML fake news detectors?</p> <p>RQ3: How does trust in the explainable ML fake news detector predict adoption intention of this AI application?</p>	<p>Users' trust levels in the software were influenced by both individuals' inherent characteristics and their perceptions of the AI application. Users' adoption intention was ultimately influenced by trust in the detector, which explained a significant amount of the variance. We also found that trust levels were higher when users perceived the application to be highly competent at detecting fake news, be highly collaborative, and have more power in working autonomously. Our findings indicate that trust is a focal element in determining users' behavioural intentions.</p>
45	[88]	<p>H1a: Explanations make it more likely that individuals will delete the proposed files.</p> <p>H1b: Providing explanations increases both affective and cognitive trust ratings.</p> <p>H1c: By providing information on why the system's suggestions are valid, users can better understand the reliability of the underlying processes. This should lead to increased credibility ratings.</p> <p>H1d: Explanations directly reduce information uncertainty.</p> <p>H2a: Information uncertainty in the system's proposals mediates the effect of explanations of the system's proposals on its acceptance.</p> <p>H2b: Information uncertainty in the system's proposals mediates the effect of explanations of the system's proposals on trust.</p> <p>H2c: Information uncertainty in the system's proposals mediates the effect of explanations of the system's proposals on credibility.</p> <p>H3a: Individuals high in need for cognition have a stronger preference for thinking about the explanations, which helps them to delete irrelevant files.</p> <p>H3b: ... build trust.</p> <p>H3c: build credibility.</p> <p>H3d: ... to reduce information uncertainty.</p> <p>H4a: Conscientiousness moderates the impact of explanations on deletion of irrelevant files.</p> <p>H4b: ... building trust.</p> <p>H4c: ... credibility.</p> <p>H4d: ... on reducing information uncertainty.</p> <p>H5: Deletion is not only an action that causes digital objects to be forgotten in external memory, but may also support intentional forgetting of associated memory content.</p>	<p>Results show the importance of presenting explanations for the acceptance of deleting suggestions in all three experiments, but also point to the need for their verifiability to generate trust in the system. However, we did not find clear evidence that deleting computer files contributes to human forgetting of the related memories.</p>

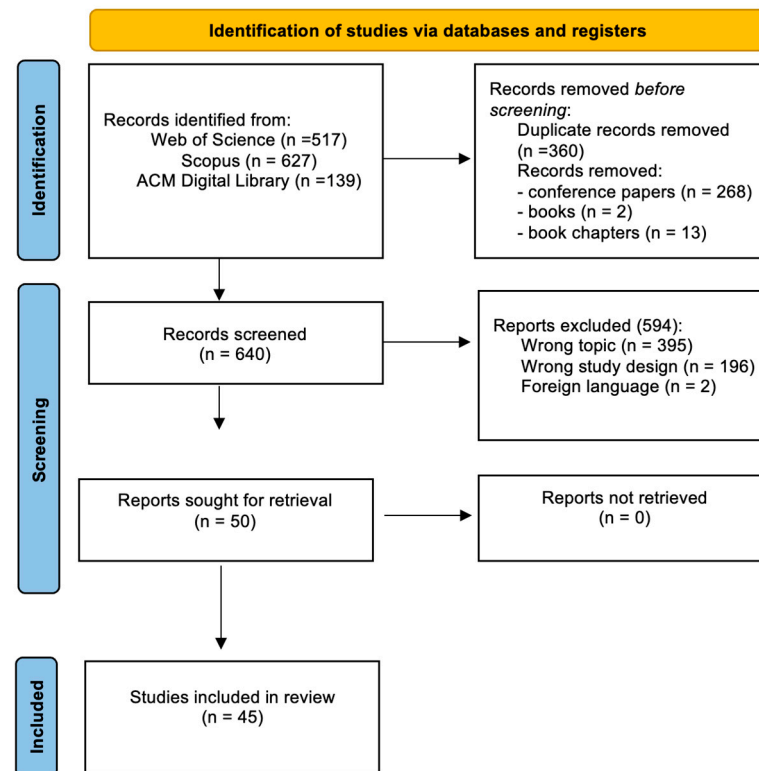


Figure 1. Study selection flow chart.

3.1. Defining and Measuring Trust

3.1.1. Definitions of Trust

The studies identified by the current review rely on different definitions of trust. Mayer et al.'s [89] definition appears common, as it was used as starting point by thirteen studies [1,46,47,49,54,55,58,63,72,76,77,80,82]. It was originally thought of as referring to human–human relationships and conceptualized 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 part”. Their model encompasses three factors of the perceived trustworthiness of the trustee that act as antecedents of trust: ability, benevolence, and integrity. Ability can be defined as having the skills and competencies to successfully complete a given task; benevolence refers to having positive intentions not based solely on self-interest; and integrity is a descriptor of the trustee’s sense of morality and justice, such that the trusted party’s behaviours are consistent, predictable, and honest. This theorization has been used to measure trust towards different applications of technology, like decision-support websites [90], recommendation agents [91], and online vendors [92].

The concept of vulnerability is also at the core of the definitions employed by three studies [74,81,87]. These studies referred to, respectively, [93]’s, [94]’s, and [9,95]’s work.

The applicability of theorizations originally conceived for human–human relationships has been widely discussed, and some argued the impossibility of applying them directly to human-to-machine trust [96]. Mcknight et al. [11], starting from the assumption that, unlike humans, technology lacks volition and moral agency, proposed three dimensions of trust in technology: functionality, or the capability of the technology; reliability, the consistency of operation; and helpfulness, indicating if the specific technology is helpful to users. This conceptualization was followed by three studies [1,55,84]. The reliability component can also be found in [4,97]’s descriptions of trust used by [75,85].

Sullivan et al. [59] referred to Lee's [98] definition of trust in technology: the attitude that an agent will help achieve an individual's goals in a situation characterized by uncertainty and vulnerability. Three studies adopted definitions of trust that directly referred to technology and AI. Yu and Li [60] referred to Höddinghaus et al. [99], who state that humans' trust in AI refers to the degree to which humans consider AI to be trustworthy. Huo et al. [64] followed Shin and Park's definition [100]: trust is the belief and intention to adopt an algorithm continuously. Finally, Cheng, Li, and Xu [65] started from Madsen and Gregor [23]: human–computer trust is the degree to which people have confidence in AI systems and are willing to take action. Instead, Zhang et al. [70] started from a conceptualization of trust in virtual assistants specifically, while Wirtz et al.'s [101] described trust as the user's confidence that the AI virtual assistant can reliably deliver a service. Ref. [88] adopted a definition referring to technical systems which states trust as the willingness to rely on them in an uncertain environment [102,103].

Lee and See [104] proposed a conceptual framework which can be considered as middle ground between human trust and technology trust and was adopted by two of the selected studies [51,61]. They suggested the use of 'trust in automation' and delineate it in three categories: performance (i.e., the operational characteristics of automation), process (i.e., the suitability of the automation for the achievement of the users' goals) and purpose (i.e., why it was designed and the designers' intent). Thirteen papers did not state a clear definition of trust [4,48,50,56,66–68,71,73,78,79,82,86]. Two studies [52,53] based their research on [105]'s general definition of trust: "a psychological state comprising the intention to accept vulnerability based upon positive expectations of the intentions or behaviour of another". Chi and Hoang Vu [57] used a general definition of trust as well, more specifically that of [106], which states that trust is the reliability of an individual ensured by one party with a given exchange relationship. As described, studies rarely refer to the same definition of trust, resulting in different ways to measure it. Some of them preferred to adapt human–human trust conceptualizations to relationships between users and AI systems. Such choice finds support in the fact that interactions between humans and between humans and technology can have similarities. This can be justified relying on the Social Response Theory (SRT; [107]), which argues that human beings are a social species and tend to attribute human-like characteristics to objects, including technological devices; as a result, they treat computers as social actors. This becomes especially relevant when talking about AI, as it is often described as capable of human qualities, such as reasoning and motivations, which also impact people's expectations and initial trust [34].

Overall, despite referring originally to human relationships, these definitions can mostly be associated with a cognitive kind of trust as they focus mainly on the ability of the trustee to achieve the expected goal. Only exception can be found in Mayer et al. [89] if looking at the antecedents of trust they define as benevolence and integrity. Though these can also be considered as attributes that are rationally considered by the trustor [14], an argument can be made that they can be characteristics that also contribute to relationship quality between trustor and trustee and take part in fostering the affective side of trust as well. Therefore, it may be noted how, even in the context of human–AI relations, the composite nature of the trust is recognized, including the interaction between the affective and cognitive components.

Other studies adopted frameworks that had already been adjusted to include artificial agents in the analysis of the relational exchange and were specifically validated in human–computer interaction settings. These seem to focus solely on functional reliability and can therefore be associated with a cognitive type of trust.

Trust Questionnaires: From Factors to Item Description

Looking at how the studies measure trust, it can be noticed that most of them focus on measuring trust itself. On the other hand, three studies [50,76,83] measured trustworthiness or perceived trustworthiness. Nevertheless, it was deemed as more relevant to go beyond the constructs and look at the items employed. The questionnaires were retrieved from the selected articles and their content analyzed. By looking at their content, each item or group of items has been synthesized with an attribute, as shown in Table 3. Furthermore, said attributes have been categorized as pertaining more to affective or cognitive trust. Attributes referring exclusively to reliability (e.g., competence, functionality) have been considered as proxies of cognitive trust. On the other hand, attributes describing characteristics that, if present, contribute to the formation of a positive and trusting relationship on a more emotional level (e.g., benevolence, care) have been regarded as associable with affective trust.

Table 3. Overview of included papers: AI systems studied, trust constructs used, and identification of trust types and attributes covered in the questionnaires.

N	Authors	AI System	Trust-Related Constructs Studies	Trust Construct	Trust Type	Attribute
1	[46]	Tool for spam review detection	Age Difference in credibility judgments Trust propensity Educational background Online shopping experience Variety of online shopping platforms usually used View on online spam reviews	<i>Trust beliefs</i>	Cognitive <hr/> Affective	Competence <hr/> Benevolence Integrity
2	[47]	Decision-making assistant	Information framing Message ownership	<i>Perceived trustworthiness</i>	Cognitive	Computer credibility
3	[48]	Interface for online insurance purchase	Perceived Usefulness Perceived Ease of Use Visibility of AI	<i>Trust in AI</i>	Cognitive <hr/> Affective	Competence Reliability Functionality Helpfulness <hr/> Benevolence Integrity
4	[49]	Text-based chatbots	Anthropomorphic attributes (perceived warmth, perceived competence, communication delay) Relationship norms (communal/exchange relationship)	<i>Trust in chatbots</i>	Cognitive <hr/> Affective	Capability <hr/> Honesty Truthfulness
5	[50]	Decision-making assistant used in public administration	Complexity of the decision process		Cognitive <hr/> Affective	Competence <hr/> Benevolence Honesty
6	[51]	Fake news detection tool	Agency locus Transparency Social presence Anthropomorphism Perceived AI's threat Social media experience	<i>Trust in judgments/system</i>	Cognitive	Correctness Accuracy Intelligence Competence
7	[52]	AI-based voice assistant systems	Information quality System quality Interaction quality	<i>Trust</i>	Cognitive	Trust
8	[53]	AI-based job application process	Social influence Personal innovativeness	<i>Perceived trust</i>	Cognitive	Perceived trust

Table 3. Cont.

N	Authors	AI System	Trust-Related Constructs Studies	Trust Construct	Trust Type	Attribute
9	[4]	Algorithm service	Transparency Accountability Fairness, explainability (FATE) Causability	Trust	Cognitive	Trust
10	[54]	Autonomous teammate (AT) in military game	Team score Trust in human teammate Trust in the team AI ethicality	Trust in the autonomous teammate	Cognitive Affective	Trust Trust
11	[55]	Smart technologies	Propensity to trust Age Education Familiarity with smart consumer technologies	Trust in AI; functionality trust in AI	Cognitive Affective	Competence Benevolence Integrity
12	[1]	Voice assistants	Perceived ease of use Perceived usefulness	Trust in the voice assistant	Cognitive Affective	Safety Competence Integrity Benevolence
13	[56]	Online content classification systems	Machine heuristic Transparency	Attitudinal trust	Cognitive Affective	Integrity Benevolence Accurateness Dependability Value Usefulness
14	[57]	AI applications in general	Interaction, empathy Anthropomorphism Communication quality COVID-19 pandemic risk	Trust in AI	Cognitive	Efficiency Safety Competence
15	[58]	AI-based voice assistant systems	Perceived usefulness Perceived ease of use Perceived enjoyment Social presence Social cognition Privacy concern	Trust	Affective	Honesty Truthfulness
16	[59]	Artificial agents in general	Perceived harm Injustice Reported wrongdoing Uncanniness	Trust in artificial agent	Cognitive	Reliability Dependability Safety
17	[60]	Decision-making assistant	Perceived effectiveness of AI Discomfort Transparency	Trust	Cognitive	Reliability
18	[61]	Medical AI	Perceived value similarity Perceived care Perceived ability Perceived uniqueness neglect	Trust in decision maker	Cognitive	Reliability
19	[62]	AI-based voice assistant systems (Siri)	Interaction, novelty value Consumer innovativeness Brand involvement Brand loyalty Perceived risk	Trust	Cognitive Affective	Competence Honesty
20	[63]	AI in companies	General trust in technology Intraorganizational trust Individual competence trust	Employees' trust in AI in the company	Cognitive Affective	Reliability Safety Competence Honesty
21	[64]	Medical AI	Self-responsibility attribution Big Five personality traits Acceptance of medical AI for independent diagnosis and treatment	Human-computer trust	Cognitive	Effectiveness Competence Reliability Helpfulness
22	[65]	Medical AI	Performance expectancy Effort expectancy Social influence	Human-computer trust	Cognitive	Effectiveness Competence Reliability Helpfulness

Table 3. Cont.

N	Authors	AI System	Trust-Related Constructs Studies	Trust Construct	Trust Type	Attribute
23	[66]	Machine learning models for automated vehicles	Interactivity of machine learning	<i>Trust in automated systems</i>	Cognitive	Security Dependability Reliability
					Affective	Integrity Familiarity
24	[67]	Medical AI	Explanation Reliability	<i>Trust in the system</i>	Cognitive	Reliability Understandability
25	[68]	Artificial neural network algorithms	Explainability	<i>Trust in AI systems</i>	Cognitive	Competence Predictability
26	[69]	Educational Artificial Intelligence Tools (EAIT)	Perceived ease of use Constructivist pedagogical beliefs Transmissive pedagogical beliefs Perceived usefulness	<i>Perceived trust</i>	Cognitive	Reliability Dependability
					Affective	Fairness
27	[70]	AI virtual assistants	Functionality (perceived usefulness, perceived ease of use) Social emotion (perceived humanity, perceived social interactivity, perceived social presence)	<i>Trust</i>	Cognitive	Favourability
					Affective	Honesty Care
28	[71]	Recommendation system	Explanations, risk Performance	<i>Trust</i>	Cognitive	Reliability Predictability Consistency Skill Capability Competency Preciseness Transparency
29	[72]	AI-based services	Digital literacy Online skills Privacy-protective behaviour	<i>Trust in AI</i>	Cognitive	Reliability Efficiency Helpfulness
30	[73]	Social robots	Age Anthropomorphism	<i>Capability Trust</i>	Cognitive	Competence
					Affective	Honesty
31	[74]	AI shopping assistant	Perceived animacy Perceived intelligence Perceived anthropomorphism	<i>Trust in Artificial Intelligence</i>	Affective	Honesty Interest respect
32	[75]	ChatGPT	Information seeking Personalization Task efficiency Playfulness Social interaction	<i>Trust</i>	Cognitive	Believability Credibility
33	[76]	AI system for candidate selection	Information about the system Trust violations and repairs	<i>Trustworthiness</i>	Cognitive	Ability
					Affective	Integrity Benevolence
34	[77]	AI	Perceived controllability Transparency	<i>Cognitive trust</i>	Cognitive	Cognitive trust
35	[78]	AI	Information seeking Concern about misinformation online	<i>Trust in AI</i>	Cognitive	Trustworthiness Rightfulness Dependability Reliability
36	[79]	AI technology	Self-efficacy Accuracy Interactivity	<i>Trust in AI</i>	Cognitive	Competence
					Affective	Honesty
37	[80]	Financial AI assistant	Autonomy	<i>Trust in automation</i>	Cognitive	Reliability
38	[81]	AI teammates	Self-efficacy Perceived anthropomorphism Perceived enjoyment Perceived rapport	<i>Trust in AI teammates</i>	Cognitive	Trustworthiness
					Affective	Honesty

Table 3. Cont.

N	Authors	AI System	Trust-Related Constructs Studies	Trust Construct	Trust Type	Attribute
39	[82]	AI chatbots	Privacy concerns Expertise Anthropomorphism Responsiveness Ease of use	<i>Trust in AI chatbots</i>	Cognitive	Trustworthiness
					Affective	Well-intendedness
40	[83]	AI recommendation system	Explanations Congruency with own judgments	<i>Perceived trustworthiness</i>	Cognitive	Competence
					Affective	Honesty Anthropomorphism
41	[84]	AI virtual assistant	Perceived risk Performance expectancy	<i>Trust</i>	Cognitive	Competence Reliability
42	[85]	AI recommendation system	System intelligence Anthropomorphism	<i>Trust</i>	Cognitive	Competence
43	[86]	AI chatbots	Empathy response Customization Interaction	<i>Trust in AI</i>	Cognitive	Competence Safety
						Dependability Competence Responsiveness Safety Reliability
44	[87]	ML fake news detectors	Performance Collaborative capacity Agency Complexity	<i>Overall AI trust, Trust in detector</i>	Cognitive	Integrity Well-intendedness Honesty
					Affective	Cognitive trust
45	[88]	Explanatory AI system	Explanations Credibility Information uncertainty	<i>Cognitive/Affective trust</i>	Cognitive	Cognitive trust
					Affective	Affective trust

It can be noticed that, apart from two studies [58,74], every research paper includes the measurement of at least one cognitive trust attribute. Conversely, affective trust attributes have been studied by twenty-four articles.

From Definition to Measurement

Comparing the approach (cognitive or affective) employed for the definition of trust and for its measurement, a few discrepancies, as well as consistencies, can be pointed out. For starters, almost every article that conceptualizes trust as a cognitive matter also measures it using attributes that recall cognitive processes. The only study that does not employ cognitive attributes is [74]. As for the studies referring to a more affective conception of trust, eight of them use both affective and cognitive attributes [1,46,49,54,58,63,76,82]. Finally, four articles measure trust with cognitive attributes despite also acknowledging the affective component in their initial definition [47,73,77,80]. To this regard, it must be said that the definition also includes a cognitive side.

Overall, the studies seem to be mostly coherent in their theoretical and measurement conceptualization of trust with respect to the distinction between cognitive and affective connotation.

3.2. Types of AI Systems

The selected papers focus on a great variety of AI systems; these will be hereby listed and described by grouping them based on their level of anthropomorphism.

Eleven of them focused on virtual assistants [1,49,52,58,62,70,74,75,82,84,86]. They considered either voice-based virtual assistants, text-based chatbots or a generic type of virtual assistant. Voice-based virtual assistants are Internet-enabled devices that provide daily technical, administrative, and social assistance to their users, including activities from setting alarms and playing music to communicating with other users [108,109]. In recent times, they also seen the implementation of natural language processing; thus, they have

become able to engage in conversational-based communication: not only can they respond to initial questions but they are also capable of asking follow-up questions [110]. It can be argued that, to different degrees based on the specific type, virtual assistants are generally characterized by a high level of anthropomorphism as they use natural language (either spoken or written) to interact directly with end users. Similar mechanisms can be found in social robots, which were object of research of only one study [73], and in autonomous teammates, as intended in [54,81].

Given their high level of anthropomorphism, these types of systems should also have higher degrees of social presence, which in turn promotes the creation of a social relationship with the user. As a consequence, the trust process is believed to also include affective elements pertaining to the bond formed between system and user. Most of the studies about highly anthropomorphic systems, apart from two, confirm this as they employ either an affective definition, affective measure components or both (as shown in Table 4). This conceptualizing and measuring set up can be deemed appropriate as people may consider these systems as social partners, similarly to how they regard human actors, therefore basing their trusting beliefs on both cognitive and affective trust. This does not intend to state the appropriateness of the anthropomorphizing behaviour but the importance of taking into account the natural human tendency to attribute human characteristics to artificial intelligence and measure trust in a way that can encompass all the nuances of this complex process.

Table 4. Cognitive/affective categorization, level of anthropomorphism and experimental stimuli.

Art N	Trust Definition: Affective Component	Trust Measurement (Aff/Cogn/Both)	System Anthropomorphism Level	Experimental Stimulus
1	Yes	Both	Low	Direct experience: task
2	Yes	Cognitive	Medium	Direct experience: task
3	No	Both	Medium	Direct experience: scenario
4	Yes	Both	High	Representation of past experience
5	No	Both	Medium	Direct experience: scenario
6	No	Cognitive	Low	Direct experience: task
7	No	Cognitive	High	Representation of past experience
8	No	Cognitive	Low	General representation
9	No	Cognitive	Low	Direct experience: task
10	Yes	Both	High	Direct experience: task
11	Yes	Both	Undefined	General representation
12	Yes	Both	High	General representation
13	No	Both	Undefined	Direct experience: task
14	No	Cognitive	Undefined	General representation
15	Yes	Affective	High	Representation of past experience
16	No	Cognitive	Undefined	Representation of past experience; direct experience: scenario
17	No	Cognitive	Medium	Direct experience: scenario
18	No	Cognitive	Medium	Direct experience: scenario

Table 4. Cont.

Art N	Trust Definition: Affective Component	Trust Measurement (Aff/Cogn/Both)	System Anthropomorphism Level	Experimental Stimulus
19	No	Both	High	Representation of past experience
20	Yes	Both	Low	Representation of past experience
21	No	Cognitive	Medium	Representation of past experience
22	No	Cognitive	Medium	Representation of past experience
23	No	Both	Low	Direct experience: task
24	No	Cognitive	Medium	Direct experience: task
25	No	Cognitive	Low	Direct experience: task
26	No	Both	Low	Representation of past experience
27	No	Both	High	Representation of past experience
28	Yes	Cognitive	Medium	Direct experience: task
29	No	Cognitive	Undefined	Representation of past experience
30	No	Both	High	Direct experience: scenario
31	No	Affective	High	Representation of past experience
32	Yes	Cognitive	High	Representation of past experience
33	Yes	Both	Medium	Direct experience: task
34	No	Cognitive	Undefined	General representation
35	No	Cognitive	Undefined	General representation
36	Yes	Both	Undefined	General representation
37	No	Cognitive	Medium	Direct experience: scenario
38	Yes	Both	High	General representation
39	No	Both	High	Representation of past experience
40	No	Both	Medium	Direct experience: scenario
41	No	Cognitive	High	General representation
42	No	Cognitive	Medium	Direct experience: scenario
43	No	Cognitive	High	General representation
44	No	Both	Low	Direct experience: task
45	No	Both	Low	Direct experience: task

Nine papers studied trust in decision-making assistants [47,50,60,71,76,80,83,85,88]. These systems were described as agents who analyze a problem they are presented with and give feedback on the best decision based on the elements they were given. They usually employ natural language to deliver the output to the user but do not necessarily engage in conversation with them; therefore, it can be hypothesized that they can elicit a medium level of anthropomorphism, lower if compared to other systems such as VAs.

It can be noticed how these studies employed, for the most part, only variables pertaining to the cognitive component of trust. As these systems are not thought to elicit relevant levels of social presence, it appears appropriate that they would mainly lead to cognitive trust, them being perceived mostly as tools rather than agents to interact and build relationships with.

Classification tools for online content were studied in four papers. More specifically, ref. [46] studied trust in spam review detection tools, refs. [51,87] focused on tools to identify fake news, and [56] focused on a system that spots online posts containing hate speech or suicidal ideation. Ref. [63] asked employees about their attitudes towards AI applications being introduced in their companies and in such positions that they would have been directly in contact with them. Ref. [66] investigated the trustworthiness of machine learning models for automated vehicles. Ref. [68] asked experts to rate the trustworthiness of some artificial neural network algorithms. Lastly, ref. [69] focused on Educational Artificial Intelligence Tools (EAIT). All these types of systems have in common a low level of anthropomorphism, that, similarly to the previous group, could lead users to perceive them more as tools rather than artificial agents subject to social norms. Interestingly, in about half the studies investigating these systems, trust was measured including both cognitive and affective components. On the contrary, it would have been expected of them to employ mostly cognitive variables, as people should not really be prone to forming socially significant relationships with systems that have little to no resemblance to humans.

Eight articles investigated a broader concept of AI system [1,4,57,59,72,77–79]. For example, ref. [1] phrased it as “smart technology”, including smart home devices (e.g., Google Nest, Ring, Blink), smart speakers (e.g., Amazon Echo, Google Home, Apple Homepod, Sonos), virtual assistants (e.g., Siri, Alexa, Cortana), and wearable devices (e.g., Fitbit, Apple Watch). In these cases, as the authors did not provide a unique target when referring to AI, it can be imagined that the participants could vary greatly in thinking about certain types of systems. Such range does not allow one to pinpoint a unique set of characteristics pertaining to the level of anthropomorphism or social presence of the systems. This variation is relevant because it may affect how trust is formed and measured, making it more difficult to interpret and compare trust constructs and attributes across studies.

Among these studies, most of them focused on cognitive components. Since they did not provide a univocal description of an AI system, this choice could have been effective in preventing participants to be drawn to think of a certain kind of system. For instance, questions pertaining to affective matters could have led them to recall attitudes and experiences related to systems with which they had a chance to build a more “personal” relationship. To this regard, cognitive aspects may have contributed to maintaining a more neutral approach and kept participants’ ideas open to a wider range of AI systems.

Overall, the studies and their differences succeed in showing how choosing which conceptualization and measure of trust should also be guided by the specific characteristics of the system that the research wants to focus on. In enacting this, special attention should be given to taking into account the mental and social processes that they elicit in users.

3.3. Experimental Stimuli

The selected papers employed different experimental stimuli before the questionnaires (Table 5). The stimuli constitute a starting point to activate participants’ cognitive processes and then be able to measure trust toward the specific AI system involved in the stimulus. Examples for each type of stimulus can be found in Table 5.

Table 5. Description of experimental stimuli used in the studies.

Type of Stimulus	Example	From
Representation of past experience	Participants were required to recall the experience of interaction with chatbots in e-commerce platforms and accordingly answer the following questions.	[49]

Table 5. Cont.

Type of Stimulus	Example	From
Direct experience: task	Participants were required to complete the Desert Survival Problem (DSP; [111]) receiving suggestions from an AI system.	[47]
Direct experience: scenario	Participants were asked to read the following vignette: “The Internal Revenue Service (IRS) is piloting a new tax return processing method to more effectively check returns for missing information. It is a simple method that is carried out by an AI. The AI scans newly submitted returns and checks them for incomplete lines or empty sections (e.g., if all the relevant boxes were ticked, all required appendices up-loaded, etc.). Following the check, the AI highlights the sections in the return where information is missing. The AI then sends an email to individuals asking them to correct their return. Individuals must comply with the AI’s request or risk facing a fine.”	[50]
General representation	Before completing the questionnaires, participants were asked whether they had experienced “AI devices”. If the answer was yes, they would go on to fill in the questionnaire. Otherwise, respondents would read an introductory brief about AI and AI applications to make sure that they understood the technical terms.	[57]

Fourteen of them measured trust, asking to participants to recall their past experiences using specific AI systems [49,52,58,59,62–65,69,70,72,74,75,82]. In these studies, participants were asked to think about past or current experiences of using a specific system, experiences that were then referenced in questions about trust. Since these are systems with which the user must have or have had an ongoing experience, it can be assumed that there has been time to create a structured relationship with them, determined by numerous occasions of interaction. As a consequence, one would expect the inclusion of affective aspects in the measurement of trust in these systems. This is for the most part reflected in the studies reviewed. In fact, among those that use recall of past or current experiences, the majority start from definitions that include the affective side and/or measure trust with even elements referable to the affective component.

Thirteen studies presented participants with a practical task that involved the use of an AI system [4,46,47,51,54,56,66–68,71,76,87,88]. These required the participant to complete a short task using an AI system, thus creating a situated user experience on which to base the trust assessment. Additionally, nine studies included written scenarios describing a process that involved the use of AI and asked participants to imagine being the protagonist [48,50,59–61,73,80,83,85]. This type of stimulus allowed for the creation of an interactive experience similar to that of the previous one, but not involving a concrete use but a figurative one, through the reading of the vignettes.

For both types of “direct” experience, therefore, participants can experience firsthand the activation of their trust processes when placed concretely in front of an AI system. Unlike past or present experience outside the experimental context, however, they have limited time to form their attitudes. For this reason, it is assumed that it is the cognitive component that may prevail in this case, as the basis for the creation of a more meaningful relationship from a social interaction point of view is not present. On the other hand, the concrete or figurative use of some systems may recall past use of similar systems. Therefore,

it may also be informative to include measurement of the affective component of trust. The studies analyzed here confirm this hypothesis: about half employ cognitive items to measure trust, while the other half use both types.

Finally, in nine studies participants were required to elicit their general representation of AI systems [1,55,57,64,77–79,84,86]. In this case, it is difficult to make assumptions about the most effective way of detecting trust, since in most cases no reference is made to a specific type of system (similarly to what has been said regarding the level of anthropomorphism). Moreover, even in the presence of one or more well-defined systems in the questionnaire delivery, it is difficult to opt for a more cognitive or affective connotation without having information about people's prior experience with them. Consequently, one cannot know whether the participants have had a chance to develop a structured relationship and attitudes toward the types of artificial intelligence and thus also a more affectively connoted trust.

3.4. Main Elements Related to Trust in AI

As previously shown in Table 3, trust was studied in relation to many other variables.

Following the discourse on affective and cognitive trust, we identified among these variables the ones that are regarded to have a more significant role in shaping one side or the other. Firstly, the variables linkable to cognitive trust will be presented, followed by the ones pertaining more to affective trust.

3.4.1. Cognitive Trust: Characteristics of the Decisional Process of AI

Performance information, explanations and transparency: A system that tells users the rationales for its decisions can be described as transparent. Transparency, presence of explanation and performance information were investigated by eleven studies [4,47,51,56,60,71,76,77,83,84,87]. Explanations about algorithmic models have been proven to enhance users' trust in the system [111–114]. It can be argued that being provided with information about how a system works directly affects the cognitive side of trust. In fact, when given information, users will rationally elaborate it and consequently decide about adopting AI. Some studies have also shown how providing information is not always helpful. For instance, ref. [112] showed how too much information negatively influences trust; ref. [115] found how users can have more trust and insight into AI when there is no explanation about how the algorithm works, rather than when there is. Moreover, studies show how trust lowers when the information given indicates that the certainty of the suggestion is low [116] or that the algorithm has limitations [117]. Additionally, given the complexity of the underlying algorithms of AI, such that experts can also find them challenging to understand, it is reasonable to consider how laypeople could not understand them [118]. Thus, it is crucial to pay attention to the way information is provided. In addition to the level of complexity of the information provided, its framing is an element to be considered. The framing effect is a cognitive bias by which people make choices depending on whether the options are presented positively or negatively [119]. Framing can assume different forms and three main types have been distinguished by [120]: risky choice, attribute and goal framing. The most relevant to the present discourse is attribute framing, which consists of an influence on how people appraise an item due to the employment of different expressions of attributes or characteristics of certain events or objects. As a result, such characteristics are manipulated depending on the context [121] and the favourability of accepting an object or event is evaluated depending on whether the message is positively or negatively framed [121–123]. The same can be applied in the case of information about AI performance and its accuracy or erroneousness that is provided to the user. For example, people may trust a system which has a "30% accuracy rate" more than one which has a

“70% error rate”. That is because, even if the accuracy rate is 30% in both cases, in the first one people could focus on the information being positively framed rather than on the actual percentages. Moreover, this effect is of crucial importance as it influences not only the framed property but also those that are not related to it [124]. The mechanism through which framing works is related to memory and emotions: positive framing is favoured because information conveyed via positive labelling highlights favourable associations in memory. As a result, people would be expected to show a more favourable attitude and behaviour toward AI when its performance is presented positively rather than negatively [121]. Nevertheless, ref. [47]’s experimental results showed that higher levels of trust in AI were more likely to be perceived by the participants that were given no information about AI’s performance, rather than those who received information, regardless of how information was framed.

Locus of agency or agency locus: Uncertainty is one of the most important factors that can lead people not to trust decisions made by AI [98]. Uncertainty in AI can be defined as the lack of the knowledge necessary to predict and explain its decisions [125]. As uncertainty reduction is a facilitator of goal achievement both in human–human and human–machine interaction, it is fundamental to trust development [126]. Only the study [51] delved into this construct. As stated by the Uncertainty Reduction Theory (URT), uncertainty has a substantial role especially when two strangers meet for the first time, because in this initial stage of interaction they have little information about each other’s attitudes, beliefs, and qualities and they are not able to accurately predict or explain each other’s behaviour [127]. The same theory also posits that individuals are naturally motivated to reduce uncertainty in order to make better plans to achieve their goals during communication. To do so, they adopt strategies to gather information that allows them to better predict and explain each other. This theory was first referred to as human–human interaction, but also has an important role in understanding the development of trust towards machines. It has been shown how predictive and explanatory knowledge about an automated system enhances user trust [126]. One of the psychological mechanisms that can impact uncertainty and, therefore, trust in AI is the so-called agency locus. Agency denotes the exercise of the capacity to act, and, with a narrow definition, it only counts intentional actions [128]. Therefore, machines cannot have it as they lack intentionality [129]. Nevertheless, it can be referred to as apparent agency, or the exercise of thinking and acting capacities that they appear to have during the interaction with humans [130]. The locus of a machine’s agency can be defined as its rules, or the cause of its apparent agency, which can be external, as created by humans, or internal, generated by the machine. This distinction has become relevant with the development of AI technologies and the massive use of machine learning techniques, which are substantially different from systems programmed by humans, because they have the capacity to define or modify decision-making rules autonomously [131]. This grants the system a certain degree of autonomy which makes it less subject to human determination. Therefore, the external agency locus pertains to systems programmed by humans and follows human-made rules that reflect human agency, while the internal agency locus characterizes AI that uses machine learning. When faced with a non-human agent, according to the three-factor theory of anthropomorphism [132], in order to reduce uncertainty about it, people are motivated to see its humanness and experience social presence. Social presence is described as a psychological state in which individuals experience a para-authentic object or an artificial object as human-like [104]. The perception of machines as human-like has been found to be helpful in reducing uncertainty as it facilitates simulation of human intelligence, a familiar type of intelligence to human users. Consequently, if people perceive the AI system as human-like, they could apply anthropocentric knowledge about how a typical human makes decisions

to understand and predict the AI system's decision-making. As a result, uncertainty is reduced and trust increases. Going back to the agency locus, if the rules are human-made, they can function as anthropomorphic cues and trigger social presence, while self-made rules are fundamentally different from human logic, and machine agency could result in less social presence, higher uncertainty and lower trust. Ref. [51] confirmed these hypotheses, finding that machine agency induced less social presence than human agency, consistently with previous similar research [133,134]. The study also showed how social presence was associated with lower uncertainty and higher trust. During the interaction, knowing that the system was programmed to follow human-made rules seemed to allow people to simulate its mental states and lead them to feel more knowledgeable about the decision-making process, similarly to what research has found about anthropomorphic attributes [135,136]. It can be noticed how, similarly to what has been said with regard to information about the system, agency locus and uncertainty perception involve mental processes that pertain more to strict rationality. Therefore, their role in shaping trust can be intended as related to the cognitive side.

Overall, all the studies that analyzed the role of variables like the ones described above also employed cognitive attributes to measure trust. Therefore, it can be stated that they were coherent in focusing on measuring both trust in a cognitive sense and other factors that could have a role in shaping it.

3.4.2. Affective Trust: Characteristics of AI's Behaviour

Social presence and social cognition: Ref. [70] studied the role played by social presence in the development of trust, as did [58], but with the addition of social cognition. Social presence is the degree of salience the other person holds during an interaction [137]. When applied to automation it has been described as the extent to which technology makes customers feel the presence of another entity [138]. Social cognition refers to how people process, store, and apply information about other people [139]. The first has been shown to have an influence on different aspects related to the interaction with technology, such as users' attitudes [140], loyalty [141], online behaviours [142,143], and trust building [2,144–146]. Moreover, research shows how technologies conveying a greater sense of social presence, such as live chat services [147], can enhance consumer trust and subsequent behaviour [145,148–150]. To achieve a better understanding of the perception of AI systems as social entities, researchers have been recently referring to the Parasocial Relationship Theory (PSRT; [151]). Such theory has been used to explain relationships between viewers/consumers and characters/celebrities. In addition, when the AI system's characteristics make it so that the interaction is human-like, it can explain how users develop a degree of closeness and intimacy with the artificial agent [26,152]. Studies have shown how having human-like features makes artificial agents being perceived as entities with mental states [26,29] more socially present and therefore subject to social norms [108] and activates social schemas that would usually be associated with human–human interactions [24]. Studies have also found that social presence, as well as social cognition, has a significant influence on trust building toward human–technology interaction [58,153]. Overall, social presence allows users to feel a somehow deep connection with a system, which in turn can foster or hinder affective trust.

Anthropomorphism: Seven studies tried to explain the influence that anthropomorphic features can have on trust development [49,57,64,73,74,82,85]. Anthropomorphism is defined as the attribution of human-like characteristics, behaviours or mental states to non-human entities such as objects, brands, animals and, more recently, technological devices [28,29]. Ref. [132] hypothesized that anthropomorphism is an inductive inference that starts from information that is readily available about humans and the person themselves

and is triggered by two motivators: effectance and sociality. Effectance can be defined as the need to interact effectively with one's environment [151], while sociality is the need and desire to establish social connections with other humans. Therefore, to have a good performance within the environment people try to reduce their uncertainty about artificial agents by attributing human characteristics to them. In addition, the lack of humans to have a social connection with can make them turn to other agents, creating a social interaction similar to the one they would have with a fellow human. Indeed, anthropomorphism is not limited to human-like appearance but can also include mental states, such as attractiveness in reasoning, moral judgements, forming intentions, and experiencing emotions [154–156]. It has been found that anthropomorphism largely influences the effectiveness of AI artifacts, as it affects the user experience [157]. Consequently, anthropomorphism also has an impact on trust. Nevertheless, this effect may also depend on the different relationship norms consumers have [158–160]. Two main types of desired relationship can guide the interaction: communal relationships and exchange relationships. The first is based on social aspects, such as caring for a partner's needs [161]. The latter is driven by the mutual exchange of comparable benefits. Therefore, people may concentrate on different anthropomorphic characteristics based on the type of relationship they are seeking. Ref. [162] identified three components of a robot's humanness: perceived physical human-likeness, perceived warmth, and perceived competence. Ref. [163] studied trust in chatbots and found those as antecedents, showing a significant effect of anthropomorphism on trust. Specifically, they found that when consumers infer chatbots as having higher warmth or higher competence based on interactions with them, they tend to develop higher levels of trust. Finally, anthropomorphism and the perception of a system as human-like can influence the relationship with the user. This can determine an effect on affective trust. For example, it can be hypothesized that a more anthropomorphic system can trigger the formation of a stronger social bond with the user, resulting in a more impactful role of the affective side of trust, either facilitating it or acting as an obstacle.

Similarly to what was said about cognitive trust, most of the studies that investigated "affective" variables such as the ones just described also measured trust including affective attributes. It is deemed significant to provide a comprehensive picture by focusing on one side of trust and on factors that can have an impact on it. On the other hand, the few studies that measured variables that can play a role in forming affective trust but only assessed cognitive trust could have missed some valuable insight on the trust development processes.

4. Discussion

A systematic review of the literature has been carried out with the main aim of exploring and critically analyzing quantitative research on trust in AI through the lenses of affective and cognitive trust, thus highlighting potential strengths and limitations. Forty-five (45) papers have been selected and examined. Results outline a very diverse scenario, as researchers describe the use of various theoretical backgrounds, experimental stimuli, measurements and AI systems. On the other hand, studies show an overall coherence in choosing methodological tools that are either cognitively or affectively oriented or that include both components.

4.1. Definitions of Trust

The results of this review highlight significant variation in how trust is conceptualized across studies, particularly in the context of human-machine interaction. Many of the studies on human-AI interaction draw on Mayer et al.'s [89] foundational framework, originally designed for human-human relationships, although its adaptation to technology-

oriented scenarios reveals both strengths and limitations. The three antecedents of this framework (ability, benevolence, and integrity) are relatively easy to map onto human–computer trust, especially the ability dimension, which aligns well with assessments of functionality and reliability that can refer to a cognitive assessment of trust. However, benevolence and integrity, which presuppose moral agency and intent, become more abstract when applied to AI systems or automated processes that lack inherent volition. Nevertheless, the inclusion of more human-like attributes such as these to the assessment of trust in AI effectively shows the close interplay between cognitive and affective trust.

Alternative approaches, such as McKnight et al.'s [90] model, attempt to bridge this gap between the conceptualization of relationship between humans and between humans and technological systems. As a matter of fact, this theory goes beyond the concept of trust involving moral and intentionality attributions and, at the same time, does not reduce trust to a mere consequence of an assessment of technical capacity to achieve a goal. They focus on functionality, reliability, and helpfulness of the system. These dimensions are particularly suitable for evaluating trust in technology, as they focus on observable performance characteristics without relying on moral assumptions. Similarly, Lee and See's [104] trust-in-automation framework, which includes performance, process, and purpose, provides a structured approach to understanding trust in automation while maintaining a clear distinction from interpersonal trust dynamics.

Nevertheless, the reliance on human–human trust theories reflects a broader trend rooted in Social Response Theory (SRT; [107]), which posits that humans attribute social characteristics to technological agents. This anthropomorphic bias may justify the use of relational concepts such as benevolence in some studies, as users often interact with AI systems as if they are social entities. This perspective is particularly relevant in contexts involving AI systems designed to mimic human reasoning or communication, as such systems may evoke expectations and emotional responses similar to those in interpersonal interactions [34].

4.2. Measuring Trust: Affective and Cognitive Components

The distinction between cognitive and affective trust emerges as a critical lens for interpreting both definitions and measurements of trust. Cognitive trust, which emphasizes rational evaluations of competence, reliability, and predictability, dominates the theoretical and methodological approaches identified in this review. This focus aligns with the functional nature of technology, where users primarily evaluate whether a system can fulfil its intended purpose. Such emphasis is evident in the widespread use of cognitive attributes in trust questionnaires, including competence, reliability, and functionality. However, affective trust, rooted in emotional connections and relational qualities, is also represented, though to a lesser extent. Attributes such as benevolence and care, discussed above, appear in several studies, suggesting a recognition of the role of emotional factors in fostering trust, particularly in contexts where users form long-term relationships with AI systems. This dual approach reflects an evolving understanding of trust as a multifaceted construct that combines rational and emotional elements.

Consequently, some discrepancies arise between how trust is defined and measured. For example, four studies that incorporate affective elements in their definitions rely solely on cognitive attributes in their measurements. This incongruence may stem from methodological challenges in operationalizing the affective side of trust, as emotional dimensions are inherently more subjective and difficult to quantify compared to cognitive aspects. In addition, let us consider that “affective” as intended in the relationship between humans may entail qualitative differences when it applies to the relationship with AI.

Indeed, there is a lack of measures that can effectively capture the complexity of both affective and cognitive trust components in AI. Hence, research should address this issue by developing tailored measurement tools that capture both components of trust, particularly when referring to AI. This is also important in light of how diverse systems can elicit different components of trust according to their specific characteristics and functions.

4.3. Types of AI Systems

The variety of AI systems explored in the reviewed studies highlights the relationship between system characteristics and user trust. Central to this analysis is the role of anthropomorphism, which influences how users perceive and interact with AI. Systems that exhibit human-like traits, whether through language, behaviour, or social cues, are deemed as more likely to evoke affective trust.

Highly anthropomorphic systems such as virtual assistants, social robots, and autonomous teammates were found to foster social sense. This aligns with prior research suggesting that systems mimicking human communication styles naturally encourage users to anthropomorphize them [110]. For example, virtual assistants equipped with conversational capabilities can elicit responses similar to interpersonal relationships. This tendency underscores why most studies on these systems incorporate the measurement of affective trust components, as the emotional bond formed with such systems mirrors the trust dynamics observed in human–human interactions. However, this anthropomorphizing behaviour, while fostering trust, also raises questions about whether users overestimate the actual capabilities and reliability of these systems. This represents an important focus for research that must take into account the ethical aspects of introducing AI systems into people’s everyday lives.

In contrast, moderately anthropomorphic systems, such as decision-making assistants, occupy a middle ground. These systems may use natural language to convey recommendations but lack sustained interaction or overt social behaviours. As a result, users perceive them more as tools than as partners, prioritizing cognitive trust over affective trust. This functional perception is consistent with the findings, where studies predominantly focused on cognitive trust measures. It raises an interesting discussion on whether this functional framing inherently limits the trust users place in such systems, especially for high-stakes decisions.

Low-anthropomorphism systems like classification tools and educational AI demonstrate minimal social cues and often do not engage users interactively. Users typically evaluate these systems based on their utility and accuracy, which naturally emphasizes cognitive trust. Yet some studies measured both cognitive and affective trust, which seems at odds with the characteristics of the systems. This discrepancy might reflect users’ general attitudes toward AI, influenced by prior interactions with other, more anthropomorphic systems, suggesting spillover effects in trust formation.

Lastly, studies exploring general AI systems without focusing on specific types highlight the complexities of trust measurement. The absence of a defined target likely pushes participants to rely on general impressions of AI, which can be influenced by numerous factors such as media narratives, personal experiences, or societal perceptions. The preference for cognitive trust measures in these studies appears conservative, as it avoids introducing biases linked to specific AI systems. However, this approach might also overlook important aspects about how trust varies across different AI contexts.

These findings collectively underscore the importance of aligning trust measures with the anthropomorphic characteristics of the AI system under study. This could prevent overestimating or underestimating the role of affective trust in user interactions, potentially skewing the conclusions drawn about user–system dynamics.

4.4. Experimental Stimuli

The experimental stimuli employed across the reviewed studies further illustrate how trust is influenced by the context in which users engage with AI systems. Stimuli such as recall-based methods, practical tasks, and scenario-based prompts each offer unique insights into trust dynamics but also carry inherent limitations.

Recall-based methods, where participants reflect on past or ongoing experiences with AI systems, provide a rich context for exploring trust. These experiences often involve repeated interactions, allowing for the formation of both cognitive and affective trust. Unsurprisingly, studies using this method frequently incorporated affective trust measures, as these systems often become integrated into users' daily lives, fostering emotional connections. However, reliance on recall introduces variability: participants may selectively remember positive or negative experiences, polarizing their responses.

Practical tasks, on the other hand, create immediate, situated interactions with AI systems, simulating real-time trust dynamics. These tasks are particularly useful for capturing cognitive trust, as participants evaluate the performance and reliability of the system in a controlled setting. The temporal constraints of such tasks, however, limit the development of affective trust, as there is insufficient time to establish a meaningful user–system relationship. Interestingly, some studies accounted for this limitation by including affective trust measures, recognizing that users might draw on memories of similar systems to inform their trust judgments. These measures can therefore allow a more complex overview and more complete information to be gathered on the relationship built between the user and the system.

Scenario-based prompts offer an alternative by immersing participants in hypothetical narratives involving AI. This method combines the immediacy of practical tasks with the flexibility to explore systems that participants may not have used directly. However, as with practical tasks, the absence of sustained interaction likely prioritizes cognitive trust. The reliance on imagination also raises concerns about ecological validity as participants' hypothetical trust judgments may not reflect how they would interact with the system in real life.

Finally, studies exploring general AI representations highlight the challenges of eliciting trust judgments without a clear target. The lack of specificity can dilute participants' responses, as they may draw on a mix of experiences, and perceptions of AI. This approach risks oversimplifying trust dynamics by neglecting the contextual factors that shape user attitudes.

Overall, the choice of experimental stimuli plays a pivotal role in shaping trust assessments. Methods emphasizing real-world experiences naturally incorporate affective trust, while those based on controlled or hypothetical interactions lean toward cognitive trust. Balancing these approaches, and recognizing their limitations, will be crucial for advancing our understanding of how trust in AI systems evolves across different contexts.

4.5. Cognitive and Affective Trust Factors

The results highlight how cognitive and affective factors shape trust in AI systems. These findings underscore the complexity of trust, suggesting that both rational and emotional dimensions are critical for fostering confidence in AI technologies.

Cognitive trust is shaped by users' logical evaluation of characteristics of AI systems. Transparency, explanations, and performance information play a pivotal role, as they allow users to rationalize AI decisions. However, the methodologies used are quite heterogeneous, both in terms of the manipulation of information and measurement and thus trust. While providing explanations enhances trust in some contexts, overly complex or negatively framed information can reduce trust [112,115].

The concept of agency locus also contributes to cognitive trust. AI systems programmed with human-made rules elicit higher social presence and reduced uncertainty compared to those employing self-generated, machine learning-driven rules [34]. This may be because human-made rules are explicitly shared with users (and are therefore known), while machine-generated rules often seem more opaque. These findings suggest that emphasizing the human influence on AI decisions could alleviate concerns about autonomy and unpredictability, in line with the idea that transparency plays a crucial role in enhancing trust.

On the other hand, affective trust emerges through users' emotional engagement with AI systems. The constructs of social presence and anthropomorphism are key drivers. Technologies that create a sense of human-like interaction, through voice, appearance, or behaviour, foster deeper affective connections and higher trust [58,70]. Anthropomorphic features such as perceived warmth and competence are particularly influential, as they align with users' expectations for social relationships [162,163]. However, the effectiveness of these dimensions depends on user preferences. For example, users seeking transactional interactions may prioritize competence over warmth, while those desiring communal connections may value empathetic behaviours more [158,159].

In sum, designing AI systems to adapt anthropomorphic features based on context and user needs could generally optimize their impact on trust. For example, balancing between providing sufficient detail and avoiding information overload, adapting explanations to user expertise could foster cognitive trust. Regarding affective trust, an effective way could be to incorporate appropriate anthropomorphic traits to strengthen affective trust while maintaining transparency about AI capabilities and preserving design ethicality. By integrating cognitive clarity with affective resonance, AI systems can build trust that is both rationally justified and emotionally compelling, paving the way for greater user acceptance and satisfaction.

Finally, an important aspect that deserves further attention is the role of cultural differences in shaping trust in human–AI interaction. The studies included in this review were conducted in different geographical regions (11 from Europe, 18 from East Asia, 6 from South Asia, and 10 from America). Each has distinct cultural norms, values and expectations regarding technology and interpersonal trust. Cultural dimensions such as individualism and collectivism, uncertainty avoidance, and power distance can significantly influence both how trust is established and how it is expressed or measured. For example, researchers might place greater emphasis on affective or cognitive trust, depending on the core values and expectations of their culture. These potential cultural variations suggest that trust in AI cannot be fully understood without considering the socio-cultural context in which interactions occur. Future research should therefore account for these cultural factors explicitly, either through comparative studies or by adapting trust measurement tools to better reflect local conceptions of trust and technology.

5. Conclusions

This review highlights the need to adapt trust models to the specific nature of human–AI interaction. Traditional human-to-human frameworks offer useful starting points but are insufficient when applied to AI systems, which are simultaneously human-like and machine-like. Therefore, trust in AI must be conceptualized through both cognitive (e.g., competence, reliability) and affective (e.g., emotional connection) lenses, while recognizing the unique, non-human qualities of these systems. Future research should aim to clearly define what constitutes trust in AI contexts and to develop specific, context-sensitive measures. A key recommendation is to treat cognitive and affective trust as distinct components, at least in early-stage investigations, to avoid conflating human psychological

constructs with machine behaviour. Constructs such as benevolence, while meaningful in human interaction, can be misleading when applied to artificial agents and should be used cautiously. The type and degree of anthropomorphism in AI influence the dominant trust path as affective trust tends to emerge more with human-like systems (e.g., social robots or virtual assistants), while cognitive trust is more relevant for functional tools (e.g., decision-support systems). This distinction should guide both empirical research and the design of AI applications that seek to promote appropriate forms of trust. To support consistency and comparability across studies, greater standardization in definitions and measurements of trust is needed. Experimental designs should be in line with the characteristics and functions of the AI systems under study, allowing for more reliable conclusions and practical implications. From a practical and policy-making perspective, these insights can inform the development of guidelines for the responsible design of AI systems, emphasizing transparency, trustworthiness, and user-centred interaction strategies. Policymakers should also consider these dimensions of trust when developing regulatory frameworks, particularly in high-risk sectors such as healthcare, education, and public services, where user trust is critical to adoption and ethical alignment. Incorporating clear, evidence-based trust criteria into design standards and certification processes will help ensure that AI systems are not only technically robust, but also socially acceptable and trustworthy.

Despite the contributions of this review, some limitations must be acknowledged. First, by focusing exclusively on quantitative methodologies, the scope of insights captured may have been constrained. Including qualitative studies in future reviews could offer important complementary perspectives, particularly in understanding the nuanced user experiences and contextual factors that influence trust formation. Qualitative approaches could provide a deeper look into the subjective dimensions of trust that quantitative methods alone may overlook.

Second, the inclusion criteria were limited to peer-reviewed journal articles, which, while ensuring the rigour and reliability of the studies reviewed, may have unintentionally excluded valuable insights from conference proceedings, book chapters, and other non-journal sources. These sources might contain innovative or emerging findings that have not yet been formally published in peer-reviewed journals.

Taken together, these limitations suggest that future reviews would benefit from a broader methodological approach and a more inclusive source selection strategy. By expanding the scope to include both qualitative and non-journal sources, future research could address existing gaps in the literature and offer a more holistic understanding of trust in human–AI interactions.

Author Contributions: Conceptualization, A.M., P.B. and L.A.; methodology, F.M., L.A. and C.D.D.; validation, A.M., D.M. and C.D.D.; formal analysis, L.A., F.M. and P.B.; investigation, L.A., P.B. and F.M.; writing—original draft preparation, L.A.; writing—review and editing, A.M., P.B. and C.D.D.; supervision, A.M. and P.B. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by: PNRR MUR project PE0000013-FAIR; EU Commission under Grant Agreement number 101094665; the research line (funds for research and publication) of the Università Cattolica del Sacro Cuore of Milan; “PON REACT EU DM 1062/21 57-I-999-1: Artificial agents, humanoid robots and human-robot interactions” funding of the Università Cattolica del Sacro Cuore of Milan.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: No data was created for this study.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Choung, H.; David, P.; Ross, A. Trust in AI and Its Role in the Acceptance of AI Technologies. *Int. J. Hum.-Comput. Interact.* **2022**, *39*, 1727–1739. [[CrossRef](#)]
2. Gefen, D.; Karahanna, E.; Straub, D.W. Trust and TAM in Online Shopping: An Integrated Model. *MIS Q.* **2003**, *27*, 51. [[CrossRef](#)]
3. Nikou, S.A.; Economides, A.A. Mobile-based assessment: Investigating the factors that influence behavioral intention to use. *Comput. Educ.* **2017**, *109*, 56–73. [[CrossRef](#)]
4. Shin, D. The effects of explainability and causability on perception, trust, and acceptance: Implications for explainable AI. *Int. J. Hum.-Comput. Stud.* **2021**, *146*, 102551. [[CrossRef](#)]
5. Shin, J.; Bulut, O.; Gierl, M.J. Development Practices of Trusted AI Systems among Canadian Data Scientists. *Int. Rev. Inf. Ethics* **2020**, *28*, 1–10. [[CrossRef](#)]
6. Bach, T.A.; Khan, A.; Hallock, H.; Beltrão, G.; Sousa, S. A Systematic Literature Review of User Trust in AI-Enabled Systems: An HCI Perspective. *Int. J. Hum.-Comput. Interact.* **2024**, *40*, 1251–1266. [[CrossRef](#)]
7. Lewis, J.D.; Weigert, A. Trust as a Social Reality. *Soc. Forces* **1985**, *63*, 967. [[CrossRef](#)]
8. Colwell, S.R.; Hogarth-Scott, S. The effect of cognitive trust on hostage relationships. *J. Serv. Mark.* **2004**, *18*, 384–394. [[CrossRef](#)]
9. McAllister, D.J. Affect-and cognition-based trust as foundations for interpersonal cooperation in organizations. *Acad. Manag. J.* **1995**, *38*, 24–59. [[CrossRef](#)]
10. Rempel, J.K.; Holmes, J.G.; Zanna, M.P. Trust in close relationships. *J. Personal. Soc. Psychol.* **1985**, *49*, 95–112. [[CrossRef](#)]
11. Mcknight, D.H.; Carter, M.; Thatcher, J.B.; Clay, P.F. Trust in a specific technology: An investigation of its components and measures. *ACM Trans. Manage. Inf. Syst.* **2011**, *2*, 1–25. [[CrossRef](#)]
12. Aurier, P.; de Lanauze, G.S. Impacts of perceived brand relationship orientation on attitudinal loyalty: An application to strong brands in the packaged goods sector. *Eur. J. Mark.* **2012**, *46*, 1602–1627. [[CrossRef](#)]
13. Johnson, D.; Grayson, K. Cognitive and affective trust in service relationships. *J. Bus. Res.* **2005**, *58*, 500–507. [[CrossRef](#)]
14. Komiak, S.Y.; Benbasat, I. The effects of personalization and familiarity on trust and adoption of recommendation agents. *MIS Q.* **2006**, *30*, 941–949. [[CrossRef](#)]
15. Moorman, C.; Zaltman, G.; Deshpande, R. Relationship between Providers and Users of Market Research: They Dynamics of Trust within & Between Organizations. *J. Mark. Res.* **1992**, *29*, 314–328. [[CrossRef](#)]
16. Luhmann, N. *Trust and Power*; John, A., Ed.; Wiley and Sons: Chichester, UK, 1979.
17. Morrow, J.L., Jr.; Hansen, M.H.; Pearson, A.W. The cognitive and affective antecedents of general trust within cooperative organizations. *J. Manag. Issues* **2004**, *16*, 48–64.
18. Chen, C.C.; Chen, X.-P.; Meindl, J.R. How Can Cooperation Be Fostered? The Cultural Effects of Individualism–Collectivism. *Acad. Manag. Rev.* **1998**, *23*, 285. [[CrossRef](#)]
19. Kim, D. Cognition-Based Versus Affect-Based Trust Determinants in E-Commerce: Cross-Cultural Comparison Study. In Proceedings of the International Conference on Information Systems, ICIS 2005, Las Vegas, NV, USA, 11–14 December 2005.
20. Dabholkar, P.A.; van Dolen, W.M.; de Ruyter, K. A dual-sequence framework for B2C relationship formation: Moderating effects of employee communication style in online group chat. *Psychol. Mark.* **2009**, *26*, 145–174. [[CrossRef](#)]
21. Ha, B.; Park, Y.; Cho, S. Suppliers' affective trust and trust in competency in buyers: Its effect on collaboration and logistics efficiency. *Int. J. Oper. Prod. Manag.* **2011**, *31*, 56–77. [[CrossRef](#)]
22. Gursoy, D.; Chi, O.H.; Lu, L.; Nunkoo, R. Consumers acceptance of artificially intelligent (AI) device use in service delivery. *Int. J. Inf. Manag.* **2019**, *49*, 157–169. [[CrossRef](#)]
23. Madsen, M.; Gregor, S. Measuring Human-Computer Trust. In Proceedings of the 11th Australasian Conference on Information Systems, Brisbane, Australia, 6–8 December 2000; Volume 53, p. 6.
24. Pennings, J.; Woiceshyn, J. *A Typology of Organizational Control and Its Metaphors: Research in the Sociology of Organizations*; JAI Press: Stamford, CT, USA, 1987.
25. Chattaraman, V.; Kwon, W.-S.; Gilbert, J.E.; Ross, K. Should AI-Based, conversational digital assistants employ social- or task-oriented interaction style? A task-competency and reciprocity perspective for older adults. *Comput. Hum. Behav.* **2019**, *90*, 315–330. [[CrossRef](#)]
26. Louie, W.-Y.G.; McColl, D.; Nejat, G. Acceptance and Attitudes Toward a Human-like Socially Assistive Robot by Older Adults. *Assist. Technol.* **2014**, *26*, 140–150. [[CrossRef](#)] [[PubMed](#)]
27. Marchetti, A.; Manzi, F.; Itakura, S.; Massaro, D. Theory of Mind and Humanoid Robots From a Lifespan Perspective. *Z. Für Psychol.* **2018**, *226*, 98–109. [[CrossRef](#)]
28. Nass, C.I.; Brave, S. Wired for speech: How voice activates and advances the human-computer relationship. In *Computer-Human Interaction*; MIT Press: Cambridge, MA, USA, 2005.
29. Golossenko, A.; Pillai, K.G.; Aroean, L. Seeing brands as humans: Development and validation of a brand anthropomorphism scale. *Int. J. Res. Mark.* **2020**, *37*, 737–755. [[CrossRef](#)]

30. Manzi, F.; Peretti, G.; Di Dio, C.; Cangelosi, A.; Itakura, S.; Kanda, T.; Ishiguro, H.; Massaro, D.; Marchetti, A. A Robot Is Not Worth Another: Exploring Children's Mental State Attribution to Different Humanoid Robots. *Front. Psychol.* **2020**, *11*, 2011. [[CrossRef](#)]
31. Gomez-Uribe, C.A.; Hunt, N. The Netflix Recommender System: Algorithms, Business Value, and Innovation. *ACM Trans. Manag. Inf. Syst.* **2016**, *6*, 1–19. [[CrossRef](#)]
32. Hengstler, M.; Enkel, E.; Duelli, S. Applied artificial intelligence and trust—The case of autonomous vehicles and medical assistance devices. *Technol. Forecast. Soc. Change* **2016**, *105*, 105–120. [[CrossRef](#)]
33. Tussyadiah, I.P.; Zach, F.J.; Wang, J. Attitudes Toward Autonomous on Demand Mobility System: The Case of Self-Driving Taxi. In *Information and Communication Technologies in Tourism*; Schegg, R., Stangl, B., Eds.; Springer International Publishing: Cham, Switzerland, 2017; pp. 755–766. [[CrossRef](#)]
34. Glikson, E.; Woolley, A.W. Human Trust in Artificial Intelligence: Review of Empirical Research. *ANNALS* **2020**, *14*, 627–660. [[CrossRef](#)]
35. Page, M.J.; McKenzie, J.E.; Bossuyt, P.M.; Boutron, I.; Hoffmann, T.C.; Mulrow, C.D.; Shamseer, L.; Tetzlaff, J.M.; Akl, E.A.; Brennan, S.E.; et al. The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ* **2021**, *372*, 71. [[CrossRef](#)]
36. Asan, O.; Bayrak, A.E.; Choudhury, A. Artificial Intelligence and Human Trust in Healthcare: Focus on Clinicians. *J. Med. Internet Res.* **2020**, *22*, e15154. [[CrossRef](#)]
37. Keysermann, M.U.; Cramer, H.; Aylett, R.; Zoll, C.; Enz, S.; Vargas, P.A. Can I Trust You? Sharing Information with Artificial Companions. In Proceedings of the 11th International Conference on Autonomous Agents and Multiagent Systems—AAMAS'12, Valencia, Spain, 4–8 June 2012; International Foundation for Autonomous Agents and Multiagent Systems: Richland, SC, USA, 2012; Volume 3, pp. 1197–1198.
38. Churchill, G.A. A Paradigm for Developing Better Measures of Marketing Constructs. *J. Mark. Res.* **1979**, *16*, 64. [[CrossRef](#)]
39. Peter, J.P. Reliability: A Review of Psychometric Basics and Recent Marketing Practices. *J. Mark. Res.* **1979**, *16*, 6. [[CrossRef](#)]
40. Churchill, G.A.; Peter, J.P. Research Design Effects on the Reliability of Rating Scales: A Meta-Analysis. *J. Mark. Res.* **1984**, *21*, 360–375. [[CrossRef](#)]
41. Kwon, H.; Trail, G. The Feasibility of Single-Item Measures in Sport Loyalty Research. *Sport Manag. Rev.* **2005**, *8*, 69–89. [[CrossRef](#)]
42. Allen, J.F.; Byron, D.K.; Dzikovska, M.; Ferguson, G.; Galescu, L.; Stent, A. Toward conversational human-computer interaction. *AI Mag.* **2001**, *22*, 27.
43. Raees, M.; Meijerink, I.; Lykourantzou, I.; Khan, V.-J.; Papangelis, K. From Explainable to Interactive AI: A Literature Review on Current Trends in Human-AI Interaction. *arXiv* **2024**, arXiv:2405.15051. [[CrossRef](#)]
44. Chu, Z.; Wang, Z.; Zhang, W. Fairness in Large Language Models: A Taxonomic Survey. *SIGKDD Explor. Newsl.* **2024**, *26*, 34–48. [[CrossRef](#)]
45. Zhang, W. AI fairness in practice: Paradigm, challenges, and prospects. *AI Mag.* **2024**, *45*, 386–395. [[CrossRef](#)]
46. Xiang, H.; Zhou, J.; Xie, B. AI tools for debunking online spam reviews? Trust of younger and older adults in AI detection criteria. *Behav. Inf. Technol.* **2022**, *42*, 478–497. [[CrossRef](#)]
47. Kim, T.; Song, H. Communicating the Limitations of AI: The Effect of Message Framing and Ownership on Trust in Artificial Intelligence. *Int. J. Hum.-Comput. Interact.* **2022**, *39*, 790–800. [[CrossRef](#)]
48. Zarifis, A.; Kawalek, P.; Azadegan, A. Evaluating If Trust and Personal Information Privacy Concerns Are Barriers to Using Health Insurance That Explicitly Utilizes AI. *J. Internet Commer.* **2021**, *20*, 66–83. [[CrossRef](#)]
49. Cheng, X.; Zhang, X.; Cohen, J.; Mou, J. Human vs. AI: Understanding the impact of anthropomorphism on consumer response to chatbots from the perspective of trust and relationship norms. *Inf. Process. Manag.* **2022**, *59*, 102940. [[CrossRef](#)]
50. Ingrams, A.; Kaufmann, W.; Jacobs, D. In AI we trust? Citizen perceptions of AI in government decision making. *Policy Internet* **2022**, *14*, 390–409. [[CrossRef](#)]
51. Liu, B. In AI We Trust? Effects of Agency Locus and Transparency on Uncertainty Reduction in Human–AI Interaction. *J. Comput.-Mediat. Commun.* **2021**, *26*, 384–402. [[CrossRef](#)]
52. Lee, O.-K.D.; Ayyagari, R.; Nasirian, F.; Ahmadian, M. Role of interaction quality and trust in use of AI-based voice-assistant systems. *J. Syst. Inf. Technol.* **2021**, *23*, 154–170. [[CrossRef](#)]
53. Kandoth, S.; Shekhar, S.K. Social influence and intention to use AI: The role of personal innovativeness and perceived trust using the parallel mediation model. *Forum Sci. Oeconomia* **2022**, *10*, 131–150. [[CrossRef](#)]
54. Schelble, B.G.; Lopez, J.; Textor, C.; Zhang, R.; McNeese, N.J.; Pak, R.; Freeman, G. Towards Ethical AI: Empirically Investigating Dimensions of AI Ethics, Trust Repair, and Performance in Human-AI Teaming. *Hum. Factors* **2022**, *66*, 1037–1055. [[CrossRef](#)]
55. Choung, H.; David, P.; Ross, A. Trust and ethics in AI. *AI Soc.* **2022**, *38*, 733–745. [[CrossRef](#)]
56. Molina, M.D.; Sundar, S.S. When AI moderates online content: Effects of human collaboration and interactive transparency on user trust. *J. Comput.-Mediat. Commun.* **2022**, *27*, zmac010. [[CrossRef](#)]

57. Chi, N.T.K.; Vu, N.H. Investigating the customer trust in artificial intelligence: The role of anthropomorphism, empathy response, and interaction. *CAAI Trans. Intell. Technol.* **2022**, *8*, 260–273. [[CrossRef](#)]
58. Pitardi, V.; Marriott, H.R. Alexa, she's not human but ... Unveiling the drivers of consumers' trust in voice-based artificial intelligence. *Psychol. Mark.* **2021**, *38*, 626–642. [[CrossRef](#)]
59. Sullivan, Y.; de Bourmont, M.; Dunaway, M. Appraisals of harms and injustice trigger an eerie feeling that decreases trust in artificial intelligence systems. *Ann. Oper. Res.* **2022**, *308*, 525–548. [[CrossRef](#)]
60. Yu, L.; Li, Y. Artificial Intelligence Decision-Making Transparency and Employees' Trust: The Parallel Multiple Mediating Effect of Effectiveness and Discomfort. *Behav. Sci.* **2022**, *12*, 127. [[CrossRef](#)]
61. Yokoi, R.; Eguchi, Y.; Fujita, T.; Nakayachi, K. Artificial Intelligence Is Trusted Less than a Doctor in Medical Treatment Decisions: Influence of Perceived Care and Value Similarity. *Int. J. Hum.-Comput. Interact.* **2021**, *37*, 981–990. [[CrossRef](#)]
62. Hasan, R.; Shams, R.; Rahman, M. Consumer trust and perceived risk for voice-controlled artificial intelligence: The case of Siri. *J. Bus. Res.* **2021**, *131*, 591–597. [[CrossRef](#)]
63. Łapińska, J.; Escher, I.; Górka, J.; Sudolska, A.; Brzustewicz, P. Employees' trust in artificial intelligence in companies: The case of energy and chemical industries in Poland. *Energies* **2021**, *14*, 1942. [[CrossRef](#)]
64. Huo, W.; Zheng, G.; Yan, J.; Sun, L.; Han, L. Interacting with medical artificial intelligence: Integrating self-responsibility attribution, human-computer trust, and personality. *Comput. Hum. Behav.* **2022**, *132*, 107253. [[CrossRef](#)]
65. Cheng, M.; Li, X.; Xu, J. Promoting Healthcare Workers' Adoption Intention of Artificial-Intelligence-Assisted Diagnosis and Treatment: The Chain Mediation of Social Influence and Human-Computer Trust. *IJERPH* **2022**, *19*, 13311. [[CrossRef](#)]
66. Gutzwiller, R.S.; Reeder, J. Dancing With Algorithms: Interaction Creates Greater Preference and Trust in Machine-Learned Behavior. *Hum. Factors* **2021**, *63*, 854–867. [[CrossRef](#)]
67. Goel, K.; Sindhgatta, R.; Kalra, S.; Goel, R.; Mutreja, P. The effect of machine learning explanations on user trust for automated diagnosis of COVID-19. *Comput. Biol. Med.* **2022**, *146*, 105587. [[CrossRef](#)] [[PubMed](#)]
68. Nakashima, H.H.; Mantovani, D.; Junior, C.M. Users' trust in black-box machine learning algorithms. *Rev. Gest.* **2022**, *31*, 237–250. [[CrossRef](#)]
69. Choi, S.; Jang, Y.; Kim, H. Influence of Pedagogical Beliefs and Perceived Trust on Teachers' Acceptance of Educational Artificial Intelligence Tools. *Int. J. Hum.-Comput. Interact.* **2022**, *39*, 910–922. [[CrossRef](#)]
70. Zhang, S.; Meng, Z.; Chen, B.; Yang, X.; Zhao, X. Motivation, Social Emotion, and the Acceptance of Artificial Intelligence Virtual Assistants-Trust-Based Mediating Effects. *Front. Psychol.* **2021**, *12*, 728495. [[CrossRef](#)]
71. De Brito Duarte, R.; Correia, F.; Arriaga, P.; Paiva, A. AI Trust: Can Explainable AI Enhance Warranted Trust? *Hum. Behav. Emerg. Technol.* **2023**, *2023*, 4637678. [[CrossRef](#)]
72. Jang, C. Coping with vulnerability: The effect of trust in AI and privacy-protective behaviour on the use of AI-based services. *Behav. Inf. Technol.* **2024**, *43*, 2388–2400. [[CrossRef](#)]
73. Jiang, C.; Guan, X.; Zhu, J.; Wang, Z.; Xie, F.; Wang, W. The future of artificial intelligence and digital development: A study of trust in social robot capabilities. *J. Exp. Theor. Artif. Intell.* **2023**, *37*, 783–795. [[CrossRef](#)]
74. Malhotra, G.; Ramalingam, M. Perceived anthropomorphism and purchase intention using artificial intelligence technology: Examining the moderated effect of trust. *JEIM* **2023**, *38*, 401–423. [[CrossRef](#)]
75. Baek, T.H.; Kim, M. Is ChatGPT scary good? How user motivations affect creepiness and trust in generative artificial intelligence. *Telemat. Inform.* **2023**, *83*, 102030. [[CrossRef](#)]
76. Langer, M.; König, C.J.; Back, C.; Hemsing, V. Trust in Artificial Intelligence: Comparing Trust Processes Between Human and Automated Trustees in Light of Unfair Bias. *J. Bus. Psychol.* **2023**, *38*, 493–508. [[CrossRef](#)]
77. Shamim, S.; Yang, Y.; Zia, N.U.; Khan, Z.; Shariq, S.M. Mechanisms of cognitive trust development in artificial intelligence among front line employees: An empirical examination from a developing economy. *J. Bus. Res.* **2023**, *167*, 114168. [[CrossRef](#)]
78. Neyazi, T.A.; Ee, T.K.; Nadaf, A.; Schroeder, R. The effect of information seeking behaviour on trust in AI in Asia: The moderating role of misinformation concern. *New Media Soc.* **2023**, *27*, 2414–2433. [[CrossRef](#)]
79. Alam, S.S.; Masukujjaman, M.; Makhbul, Z.K.M.; Ali, M.H.; Ahmad, I.; Al Mamun, A. Experience, Trust, eWOM Engagement and Usage Intention of AI Enabled Services in Hospitality and Tourism Industry: Moderating Mediating Analysis. *J. Qual. Assur. Hosp. Tour.* **2023**, *25*, 1635–1663. [[CrossRef](#)]
80. Schreiberlmayr, S.; Moradbakhti, L.; Mara, M. First impressions of a financial AI assistant: Differences between high trust and low trust users. *Front. Artif. Intell.* **2023**, *6*, 1241290. [[CrossRef](#)]
81. Hou, K.; Hou, T.; Cai, L. Exploring Trust in Human-AI Collaboration in the Context of Multiplayer Online Games. *Systems* **2023**, *11*, 217. [[CrossRef](#)]
82. Li, J.; Wu, L.; Qi, J.; Zhang, Y.; Wu, Z.; Hu, S. Determinants Affecting Consumer Trust in Communication With AI Chatbots: The Moderating Effect of Privacy Concerns. *J. Organ. End. User Comput.* **2023**, *35*, 1–24. [[CrossRef](#)]
83. Selten, F.; Robeer, M.; Grimmelhuijsen, S. 'Just like I thought': Street-level bureaucrats trust AI recommendations if they confirm their professional judgment. *Public Adm. Rev.* **2023**, *83*, 263–278. [[CrossRef](#)]

84. Xiong, Y.; Shi, Y.; Pu, Q.; Liu, N. More trust or more risk? User acceptance of artificial intelligence virtual assistant. *Hum. FTRS Erg. MFG SVC* **2024**, *34*, 190–205. [[CrossRef](#)]
85. Song, J.; Lin, H. Exploring the effect of artificial intelligence intellect on consumer decision delegation: The role of trust, task objectivity, and anthropomorphism. *J. Consum. Behav.* **2024**, *23*, 727–747. [[CrossRef](#)]
86. Nguyen, V.T.; Phong, L.T.; Chi, N.T.K. The impact of AI chatbots on customer trust: An empirical investigation in the hotel industry. *CBTH* **2023**, *18*, 293–305. [[CrossRef](#)]
87. Shin, J.; Chan-Olmsted, S. User Perceptions and Trust of Explainable Machine Learning Fake News Detectors. *Int. J. Commun.* **2023**, *17*, 518–540.
88. Göbel, K.; Niessen, C.; Seufert, S.; Schmid, U. Explanatory machine learning for justified trust in human-AI collaboration: Experiments on file deletion recommendations. *Front. Artif. Intell.* **2022**, *5*, 919534. [[CrossRef](#)]
89. Mayer, R.C.; Davis, J.H.; Schoorman, F.D. An Integrative Model of Organizational Trust. *Acad. Manag. Rev.* **1995**, *20*, 709. [[CrossRef](#)]
90. McKnight, D.H.; Choudhury, V.; Kacmar, C. Developing and Validating Trust Measures for e-Commerce: An Integrative Typology. *Inf. Syst. Res.* **2002**, *13*, 334–359. [[CrossRef](#)]
91. Wang, W.; Benbasat, I. Attributions of Trust in Decision Support Technologies: A Study of Recommendation Agents for E-Commerce. *J. Manag. Inf. Syst.* **2008**, *24*, 249–273. [[CrossRef](#)]
92. Oliveira, T.; Alinho, M.; Rita, P.; Dhillon, G. Modelling and testing consumer trust dimensions in e-commerce. *Comput. Hum. Behav.* **2017**, *71*, 153–164. [[CrossRef](#)]
93. Ozdemir, S.; Zhang, S.; Gupta, S.; Bebek, G. The effects of trust and peer influence on corporate brand—Consumer relationships and consumer loyalty. *J. Bus. Res.* **2020**, *117*, 791–805. [[CrossRef](#)]
94. Grazioli, S.; Jarvenpaa, S.L. Perils of Internet fraud: An empirical investigation of deception and trust with experienced Internet consumers. *IEEE Trans. Syst. Man. Cybern. A* **2000**, *30*, 395–410. [[CrossRef](#)]
95. Lewicki, R.J.; Tomlinson, E.C.; Gillespie, N. Models of Interpersonal Trust Development: Theoretical Approaches, Empirical Evidence, and Future Directions. *J. Manag.* **2006**, *32*, 991–1022. [[CrossRef](#)]
96. Madhavan, P.; Wiegmann, D.A. Effects of Information Source, Pedigree, and Reliability on Operator Interaction With Decision Support Systems. *Hum. Factors* **2007**, *49*, 773–785. [[CrossRef](#)]
97. Moorman, C.; Deshpandé, R.; Zaltman, G. Factors Affecting Trust in Market Research Relationships. *J. Mark.* **1993**, *57*, 81–101. [[CrossRef](#)]
98. Lee, M.K. Understanding perception of algorithmic decisions: Fairness, trust, and emotion in response to algorithmic management. *Big Data Soc.* **2018**, *5*, 205395171875668. [[CrossRef](#)]
99. Höddinghaus, M.; Sondern, D.; Hertel, G. The automation of leadership functions: Would people trust decision algorithms? *Comput. Hum. Behav.* **2021**, *116*, 106635. [[CrossRef](#)]
100. Shin, D.; Park, Y.J. Role of fairness, accountability, and transparency in algorithmic affordance. *Comput. Hum. Behav.* **2019**, *98*, 277–284. [[CrossRef](#)]
101. Wirtz, J.; Patterson, P.G.; Kunz, W.H.; Gruber, T.; Lu, V.N.; Paluch, S.; Martins, A. Brave new world: Service robots in the frontline. *JOSM* **2018**, *29*, 907–931. [[CrossRef](#)]
102. Komiak, S.X.; Benbasat, I. Understanding Customer Trust in Agent-Mediated Electronic Commerce, Web-Mediated Electronic Commerce, and Traditional Commerce. *Inf. Technol. Manag.* **2004**, *5*, 181–207. [[CrossRef](#)]
103. Meeßen, S.M.; Thielsch, M.T.; Hertel, G. Trust in Management Information Systems (MIS): A Theoretical Model. *Z. Arb.-Organ. AO* **2020**, *64*, 6–16. [[CrossRef](#)]
104. Lee, J.D.; See, K.A. Trust in Automation: Designing for Appropriate Reliance. *Hum. Factors J. Hum. Factors Ergon. Soc.* **2004**, *46*, 50–80. [[CrossRef](#)] [[PubMed](#)]
105. Rousseau, D.M.; Sitkin, S.B.; Burt, R.S.; Camerer, C. Not So Different After All: A Cross-Discipline View Of Trust. *AMR* **1998**, *23*, 393–404. [[CrossRef](#)]
106. Rotter, J.B. Interpersonal trust, trustworthiness, and gullibility. *Am. Psychol.* **1980**, *35*, 1–7. [[CrossRef](#)]
107. Nass, C.; Moon, Y. Machines and Mindlessness: Social Responses to Computers. *J. Soc. Issues* **2000**, *56*, 81–103. [[CrossRef](#)]
108. Han, S.; Yang, H. Understanding adoption of intelligent personal assistants: A parasocial relationship perspective. *IMDS* **2018**, *118*, 618–636. [[CrossRef](#)]
109. Santos, J.; Rodrigues, J.J.P.C.; Silva, B.M.C.; Casal, J.; Saleem, K.; Denisov, V. An IoT-based mobile gateway for intelligent personal assistants on mobile health environments. *J. Netw. Comput. Appl.* **2016**, *71*, 194–204. [[CrossRef](#)]
110. Hoy, M.B. Alexa, Siri, Cortana, and More: An Introduction to Voice Assistants. *Med. Ref. Serv. Q.* **2018**, *37*, 81–88. [[CrossRef](#)] [[PubMed](#)]
111. Pu, P.; Chen, L. Trust building with explanation interfaces. In Proceedings of the 11th International Conference on Intelligent User Interfaces, Sydney, Australia, 29 January–1 February 2006; pp. 93–100. [[CrossRef](#)]

112. Bunt, A.; McGrenere, J.; Conati, C. Understanding the Utility of Rationale in a Mixed-Initiative System for GUI Customization. In *User Modeling*; Lecture Notes in Computer Science; Conati, C., McCoy, K., Paliouras, G., Eds.; Springer: Berlin/Heidelberg, Germany, 2007; Volume 4511, pp. 147–156. [[CrossRef](#)]
113. Kizilcec, R.F. How Much Information?: Effects of Transparency on Trust in an Algorithmic Interface. In Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems, San Jose, CA, USA, 7–12 May 2016; pp. 2390–2395. [[CrossRef](#)]
114. Lafferty, J.C.; Eady, P.M.; Pond, A.W. *The Desert Survival Problem: A Group Decision Making Experience for Examining and Increasing Individual and Team Effectiveness: Manual*; Experimental Learning Methods: Plymouth, MI, USA, 1974.
115. Zhang, J.; Curley, S.P. Exploring Explanation Effects on Consumers' Trust in Online Recommender Agents. *Int. J. Hum.-Comput. Interact.* **2018**, *34*, 421–432. [[CrossRef](#)]
116. Papenmeier, A.; Kern, D.; Englebienne, G.; Seifert, C. It's Complicated: The Relationship between User Trust, Model Accuracy and Explanations in AI. *ACM Trans. Comput.-Hum. Interact.* **2022**, *29*, 1–33. [[CrossRef](#)]
117. Lim, B.Y.; Dey, A.K. Investigating intelligibility for uncertain context-aware applications. In Proceedings of the 13th International Conference on Ubiquitous Computing, Beijing, China, 17–21 September 2011; pp. 415–424. [[CrossRef](#)]
118. Cai, C.J.; Jongejan, J.; Holbrook, J. The effects of example-based explanations in a machine learning interface. In Proceedings of the 24th International Conference on Intelligent User Interfaces, Marina del Rey, CA, USA, 17–20 March 2019; pp. 258–262. [[CrossRef](#)]
119. Ananny, M.; Crawford, K. Seeing without knowing: Limitations of the transparency ideal and its application to algorithmic accountability. *New Media Soc.* **2018**, *20*, 973–989. [[CrossRef](#)]
120. Plous, S. The psychology of judgment and decision making. *J. Mark.* **1994**, *58*, 119.
121. Levin, I.P.; Schneider, S.L.; Gaeth, G.J. All Frames Are Not Created Equal: A Typology and Critical Analysis of Framing Effects. *Organ. Behav. Hum. Decis. Process.* **1998**, *76*, 149–188. [[CrossRef](#)] [[PubMed](#)]
122. Levin, P.; Schnittjer, S.K.; Thee, S.L. Information framing effects in social and personal decisions. *J. Exp. Soc. Psychol.* **1988**, *24*, 520–529. [[CrossRef](#)]
123. Davis, M.A.; Bobko, P. Contextual effects on escalation processes in public sector decision making. *Organ. Behav. Hum. Decis. Process.* **1986**, *37*, 121–138. [[CrossRef](#)]
124. Dunegan, K.J. Framing, cognitive modes, and image theory: Toward an understanding of a glass half full. *J. Appl. Psychol.* **1993**, *78*, 491–503. [[CrossRef](#)]
125. Wong, R.S. An Alternative Explanation for Attribute Framing and Spillover Effects in Multidimensional Supplier Evaluation and Supplier Termination: Focusing on Asymmetries in Attention. *Decis. Sci.* **2021**, *52*, 262–282. [[CrossRef](#)]
126. Berger, C.R. Communicating under uncertainty. In *Interpersonal Processes: New Directions in Communication Research*; Sage Publications, Inc.: Thousand, OK, USA, 1987; pp. 39–62.
127. Jian, J.-Y.; Bisantz, A.M.; Drury, C.G. Foundations for an Empirically Determined Scale of Trust in Automated Systems. *Int. J. Cogn. Ergon.* **2000**, *4*, 53–71. [[CrossRef](#)]
128. Berger, C.R.; Calabrese, R.J. Some explorations in initial interaction and beyond: Toward a developmental theory of interpersonal communication. *Hum. Comm. Res.* **1975**, *1*, 99–112. [[CrossRef](#)]
129. Schlosser, M.E. Dual-system theory and the role of consciousness in intentional action. In *Free Will, Causality, and Neuroscience*; Missal, M., Sims, A., Eds.; Brill: Leiden Bernard Feltz, The Netherlands, 2019.
130. Dennett, D.C. Précis of The Intentional Stance. *Behav. Brain Sci.* **1988**, *11*, 495. [[CrossRef](#)]
131. Takayama, L. Telepresence and Apparent Agency in Human–Robot Interaction. In *The Handbook of the Psychology of Communication Technology*, 1st ed.; Sundar, S.S., Ed.; Wiley: Hoboken, NJ, USA, 2015; pp. 160–175. [[CrossRef](#)]
132. Mittelstadt, B.D.; Allo, P.; Taddeo, M.; Wachter, S.; Floridi, L. The ethics of algorithms: Mapping the debate. *Big Data Soc.* **2016**, *3*, 205395171667967. [[CrossRef](#)]
133. Epley, N.; Waytz, A.; Cacioppo, J.T. On seeing human: A three-factor theory of anthropomorphism. *Psychol. Rev.* **2007**, *114*, 864–886. [[CrossRef](#)]
134. Fox, J.; Ahn, S.J.; Janssen, J.H.; Yeykelis, L.; Segovia, K.Y.; Bailenson, J.N. Avatars Versus Agents: A Meta-Analysis Quantifying the Effect of Agency on Social Influence. *Hum.-Comput. Interact.* **2015**, *30*, 401–432. [[CrossRef](#)]
135. Oh, C.S.; Bailenson, J.N.; Welch, G.F. A Systematic Review of Social Presence: Definition, Antecedents, and Implications. *Front. Robot. AI* **2018**, *5*, 114. [[CrossRef](#)]
136. Kiesler, S.; Goetz, J. Mental models of robotic assistants. In Proceedings of the CHI'02 Extended Abstracts on Human Factors in Computing Systems, Minneapolis, MN, USA, 20–25 April 2002; pp. 576–577. [[CrossRef](#)]
137. Ososky, S.; Philips, E.; Schuster, D.; Jentsch, F. A Picture is Worth a Thousand Mental Models: Evaluating Human Understanding of Robot Teammates. *Proc. Hum. Factors Ergon. Soc. Annu. Meet.* **2013**, *57*, 1298–1302. [[CrossRef](#)]
138. Short, J.; Williams, E.; Christie, B. *The Social Psychology of Telecommunications*; Wiley: New York, NY, USA, 1976.
139. Van Doorn, J.; Mende, M.; Noble, S.M.; Hulland, J.; Ostrom, A.L.; Grewal, D.; Petersen, J.A. Emergence of Automated Social Presence in Organizational Frontlines and Customers' Service Experiences. *J. Serv. Res.* **2017**, *20*, 43–58. [[CrossRef](#)]

140. Fiske, S.T.; Macrae, C.N. (Eds.) *The SAGE Handbook of Social Cognition*; SAGE: London, UK, 2012.
141. Hassanein, K.; Head, M. Manipulating perceived social presence through the web interface and its impact on attitude towards online shopping. *Int. J. Hum.-Comput. Stud.* **2007**, *65*, 689–708. [[CrossRef](#)]
142. Cyr, D.; Hassanein, K.; Head, M.; Ivanov, A. The role of social presence in establishing loyalty in e-Service environments. *Interact. Comput.* **2007**, *19*, 43–56. [[CrossRef](#)]
143. Chung, N.; Han, H.; Koo, C. Adoption of travel information in user-generated content on social media: The moderating effect of social presence. *Behav. Inf. Technol.* **2015**, *34*, 902–919. [[CrossRef](#)]
144. Ogara, S.O.; Koh, C.E.; Prybutok, V.R. Investigating factors affecting social presence and user satisfaction with Mobile Instant Messaging. *Comput. Hum. Behav.* **2014**, *36*, 453–459. [[CrossRef](#)]
145. Gefen, D.; Straub, D.W. Consumer trust in B2C e-Commerce and the importance of social presence: Experiments in e-Products and e-Services. *Omega* **2004**, *32*, 407–424. [[CrossRef](#)]
146. Lu, B.; Fan, W.; Zhou, M. Social presence, trust, and social commerce purchase intention: An empirical research. *Comput. Hum. Behav.* **2016**, *56*, 225–237. [[CrossRef](#)]
147. Ogonowski, A.; Montandon, A.; Botha, E.; Reyneke, M. Should new online stores invest in social presence elements? The effect of social presence on initial trust formation. *J. Retail. Consum. Serv.* **2014**, *21*, 482–491. [[CrossRef](#)]
148. McLean, G.; Osei-Frimpong, K.; Wilson, A.; Pitardi, V. How live chat assistants drive travel consumers' attitudes, trust and purchase intentions: The role of human touch. *IJCHM* **2020**, *32*, 1795–1812. [[CrossRef](#)]
149. Hassanein, K.; Head, M.; Ju, C. A cross-cultural comparison of the impact of Social Presence on website trust, usefulness and enjoyment. *IJEB* **2009**, *7*, 625. [[CrossRef](#)]
150. Mackey, K.R.M.; Freyberg, D.L. The Effect of Social Presence on Affective and Cognitive Learning in an International Engineering Course Taught via Distance Learning. *J. Eng. Edu* **2010**, *99*, 23–34. [[CrossRef](#)]
151. Ye, Y.; Zeng, W.; Shen, Q.; Zhang, X.; Lu, Y. The visual quality of streets: A human-centred continuous measurement based on machine learning algorithms and street view images. *Environ. Plan. B-Urban. Anal. City Sci.* **2019**, *46*, 1439–1457. [[CrossRef](#)]
152. Horton, D.; Wohl, R.R. Mass Communication and Para-Social Interaction: Observations on Intimacy at a Distance. *Psychiatry* **1956**, *19*, 215–229. [[CrossRef](#)]
153. Sproull, L.; Subramani, M.; Kiesler, S.; Walker, J.; Waters, K. When the Interface Is a Face. *Hum.-Comp. Interact.* **1996**, *11*, 97–124. [[CrossRef](#)]
154. White, R.W. Motivation reconsidered: The concept of competence. *Psychol. Rev.* **1959**, *66*, 297–333. [[CrossRef](#)] [[PubMed](#)]
155. Di Dio, C.; Manzi, F.; Peretti, G.; Cangelosi, A.; Harris, P.L.; Massaro, D.; Marchetti, A. Shall I Trust You? From Child-Robot Interaction to Trusting Relationships. *Front. Psychol.* **2020**, *11*, 469. [[CrossRef](#)]
156. Kim, S.; McGill, A.L. Gaming with Mr. Slot or Gaming the Slot Machine? Power, Anthropomorphism, and Risk Perception. *J. Consum. Res.* **2011**, *38*, 94–107. [[CrossRef](#)]
157. Morrison, M.; Lăzăroiu, G. Cognitive Internet of Medical Things, Big Healthcare Data Analytics, and Artificial intelligence-based Diagnostic Algorithms during the COVID-19 Pandemic. *Am. J. Med. Res.* **2021**, *8*, 23. [[CrossRef](#)]
158. Schanke, S.; Burtch, G.; Ray, G. Estimating the Impact of 'Humanizing' Customer Service Chatbots. *Inf. Syst. Res.* **2021**, *32*, 736–751. [[CrossRef](#)]
159. Aggarwal, P. The Effects of Brand Relationship Norms on Consumer Attitudes and Behavior. *J. Consum. Res.* **2004**, *31*, 87–101. [[CrossRef](#)]
160. Gao, Y.; Zhang, L.; Wei, W. The effect of perceived error stability, brand perception, and relationship norms on consumer reaction to data breaches. *Int. J. Hosp. Manag.* **2021**, *94*, 102802. [[CrossRef](#)]
161. Li, X.; Chan, K.W.; Kim, S. Service with Emoticons: How Customers Interpret Employee Use of Emoticons in Online Service Encounters. *J. Consum. Res.* **2019**, *45*, 973–987. [[CrossRef](#)]
162. Shuqair, S.; Pinto, D.C.; So, K.K.F.; Rita, P.; Mattila, A.S. A pathway to consumer forgiveness in the sharing economy: The role of relationship norms. *Int. J. Hosp. Manag.* **2021**, *98*, 103041. [[CrossRef](#)]
163. Belanche, D.; Casaló, L.V.; Schepers, J.; Flavián, C. Examining the effects of robots' physical appearance, warmth, and competence in frontline services: The Humanness-Value-Loyalty model. *Psychol. Mark.* **2021**, *38*, 2357–2376. [[CrossRef](#)]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.