1 2 9 0

UNIVERSIDADE D
COIMBRA

Paulo Jorge Nogueira Ramos

PSYCHOSOCIAL FACTORS ASSOCIATED WITH TRUST IN HUMAN-ROBOT INTERACTION IN AN INDUSTRIAL CONTEXT: A SYSTEMATIC REVIEW OF LITERATURE

Dissertação no âmbito do Mestrado Integrado em Psicologia, ramo de Psicologia das Organizações e do Trabalho orientada pela Professora Doutora Carla Maria Santos Carvalho e pela Professora Doutora Ana Luísa Sousa Pinto e apresentada à Faculdade de Psicologia e de Ciências da Educação da Universidade de Coimbra.

fevereiro de 2021

Psychosocial Factors Associated with Trust in Human-Robot Interaction in an Industrial Context: A Systematic Review of the Literature

Paulo Nogueira Ramos

Faculdade de Psicologia e de Ciências da Educação da Universidade de Coimbra

Psicologia das Organizações e do Trabalho

Professora Doutora Maria Luísa Pinto e Professora Doutora Carla Maria Santos Carvalho

20 de fevereiro de 2021

“*Some people choose to see the ugliness in this world. The disarray. I choose to see the beauty. To believe there is an order to our days, a purpose.*”

Westworld

Agradecimentos

A realização desta dissertação de mestrado contou com apoios dos quais não me poderia esquecer neste momento. À Professora Doutora Ana Luísa Pinto e à Professora Doutora Carla Carvalho tenho a agradecer todo o apoio e discussão na orientação que fomentou o meu sucesso nesta etapa. A dedicação que demonstraram é inegável e merece todo o meu respeito e carinho. À Professora Doutora Leonor Pais agradeço o primeiro passo dado na direção desta especialização. À Professora Doutora Lisete Mónico, por acreditar nas minhas causas. Ao Professor Doutor Armando Mónica pelas portas que se abriram na minha mente. À Professora Doutora Paula Paixão pelo incentivo para eu pensar fora da caixa. Ao Professor Doutor José Manuel Pureza pelo desassossego e pela lição de que não somos loucos se a intenção for melhorar o mundo.

Obrigado à minha recente, mas intensa equipa do CIAAC por estarem comigo nesta demanda. Aos biólogos, aos internacionalistas e aos psicólogos que marcaram a minha vida: muito obrigado! Nada disto faria sentido sem a vossa presença no meu percurso de ser biopsicossocial. José Cerca, devo-te a aprendizagem que mesmo quando parece que estamos a perder os nossos sonhos a única solução que resta é continuar a lutar por eles. Pedro Castelão Baptista, sem ti não tinha conseguido continuar este curso. Jorge Soares da Silva, a ti agradeço por me mostrares que nem sempre temos de ser sérios. Diogo Videira, um brinde às nossas diretas a trabalhar. Alexandre Ferreira, outro ao companheirismo e teimosia das nossas conversas. Obrigado ao Aires por não me fazer ser o único com incertezas de percursos, à Joana

Costa pela ansiedade empática, ao Sandro Serrano pelas discussões intelectuais e amizade e à Sara Barbosa pela sua garra e revolta. Ao Luís Silva, meu Presidente humano! Ao Tiago Gonçalves e ao Eduardo Mateus, as duas pessoas que durante anos lidam com o meu pior e me fazem acreditar que consigo sair dele da melhor forma. Obrigado! À minha família: à minha mãe, ao meu pai, à minha irmã e ao meu avô por aceitarem as minhas indecisões na minha procura de sentido. À minha tia madrinha, ao meu tio e ao meu primo agradeço a presença mesmo na distância. E, claro, ao Benjamim pelas traquinices que faz que animam sempre os meus dias. À Emília Oliveira agradeço o facto de, nos últimos 2 anos, me ter ajudado a aguentar o peso do mundo. Sem ela eu não teria resistido. O que eu conquisto é possível com a força e amor dela. Obrigado, melhor amiga e companheira... e melhor pessoa que conheço. À minha avó: muito obrigado por teres feito parte da minha vida. Eu sei que não estás cá, mas todos os dias me lembro de lutar por um mundo mais justo e isso foste tu sem saber que o ensinaste. A tua existência blindou as minhas causas.

E, finalmente... a todos os cães e gatos que passaram pela minha vida. Eles são os melhores!

Resumo

Atendendo ao desenvolvimento da robótica e à sua integração nos contextos industriais, a relação entre trabalhadores e robôs deve ser otimizada. Assim sendo, a confiança na interação humano-robot (HRI) é um tema relevante. Destarte, o objetivo deste trabalho é identificar os fatores psicossociais que impactam de fora significante a confiança na HRI nos contextos industriais e proceder ao seu mapeamento, integrando-as em dimensões que podem definir a sua natureza. Para tal, a presente investigação considerou o período de tempo entre 2015 e o início de 2021 e foi feita recorrendo às bases de dados Web of Science e SCOPUS. 36 fatores psicossociais diferentes foram encontrados em 12 artigos distintos e 16 destas variáveis foram, simultaneamente, respostas cognitivo-comportamentais e fatores sociais de acordo com a presente classificação. Cinco *proceedings*, cinco artigos e duas revisões incluíam-se nos documentos considerados, sendo nove destes respeitantes a simulações. A escassez de resultados foi considerada uma troca necessária em prol da garantia de informação atualizada, uma vez que estudos desatualizados no âmbito da tecnologia podem ter efeitos negativos na respetiva relevância. O presente estudo pode ajudar a identificar a razão por detrás da influência na confiança dos fatores considerados por, primeiramente, dar a conhecer o seu tipo psicossocial, criando a possibilidade de se desenvolverem *designs* orientados para a otimização da confiança nos trabalhadores na HRI, na Indústria 4.0.

Palavras-chave: fatores psicossociais, confiança, interação humano-robô; industrial; robô colaborativo

Nota: A presente dissertação seguiu esta formatação, na medida em que vai ser submetida para a revista *Computers in Human Behavior*.

Abstract

Taking into consideration the development of robotics and its integration in industrial contexts, the relationship between workers and robots shall be optimized. In this sense, trust in human-robot interaction (HRI) is a relevant point. Thus, the objective of this work is to identify the significant psychosocial factors that impact trust in HRI in industrial contexts and map them into dimensions that can define their nature. To this end, the research included the period from 2015 to early 2021 and was carried out using the databases Web of Science and SCOPUS. 36 different psychosocial factors were discovered in 12 different articles, and 16 of those variables were simultaneously cognitive behavioral responses and social factors according to our classification. Five proceedings, five articles and two reviews were among the documents included, with nine being considered simulations. The scarcity of results was obtained as a tradeoff to guarantee the updated information, since outdated studies concerning technology can have negative effects on their relevance. This study can help identify the reason for factors influencing trust starting by knowing their psychosocial type, creating the possibility for designs oriented to maximize trust of workers in HRI in Industry 4.0.

Keywords: psychosocial factors; trust; human-robot interaction; industrial; cobot

Highlights:

- 12 documents were found containing 36 psychosocial factors impacting trust in HRI.
- The psychosocial factors were dimensioned into mood status, cognitive behavioral responses, social factors, and others.
- 16 out of 36 psychosocial factors are simultaneously dimensioned into cognitive behavioral responses and social factors.

- Nine out of 12 of the documents found on Web of Science and SCOPUS were done in simulated environments.

Table of contents

1. Introduction

Technological advancements impact our everyday lives, not only at the individual level, but also at the group and social ones. With the exponential development of technology, it is increasingly harder to grasp and predict the consequences of innovation (Morrar et al., 2017). Thus, how to harness technological advancement towards the improvement of our quality of life is a major concern, as Morrar and colleagues (2017) point out, and a central topic on the debate about Industry 4.0. The professional dimension is not to be forgotten as it is an important level of analysis. Industry 4.0, first mentioned in Hannover Fair in 2011 (Sung, 2018) and, as defined by Ahuett-Garza and Kurfess (2018), is the integration of diverse technologies and agents with the aim of augmenting parameters of a production system such as efficiency and responsiveness. Kagermann et al. (2013) refer this industrial revolution will upgrade the level of socio-technical interaction through autonomous, self-configuring and self-controlling networks of manufacturing resources such as robots. Enhancements on manufacturing industries production can be made using Artificial Intelligence (AI) and robotics (Bednar & Welch, 2020). The partnership between human and robots is generally useful in unreachable or dangerous environments (Hancock et al., 2011) which can positively impact the safety and well-being of the worker in the industrial context if well managed.

The sociotechnical systems (STS) theory aims to enhance the performance of organizational systems through the recognition of the impact of humans on the operation of technology (Pasmore et al., 2019). STS theory is also relevant in the extent to which it reveals the need of joint integration between a social system – people doing their work – and a technical system – technical requirements for them to do their work (Fox, 1995). Prioritizing the joint optimization between human and robots in industrial contexts is crucial to achieve better results in terms of well-being, security, and productivity. The Human-Robot Interaction (HRI) is an important example of what can be the interdependence of the social (human) and technical (robot) systems. Understanding that interaction is fundamental due to the constant technological innovation flow created by Industry 4.0. It can guide the correct creation of strategies to implement robots taking into consideration variables that optimize its usage by workers. One possible and relevant way of achieving this is through the comprehension of the psychosocial factors that impact trust of workers in industrial contexts when they are interacting with robots. On this matter, different studies focused on gathering and study variables that can impact trust in HRI. One of the possible categorizations of factors made in literature is the division into three categories: human-related, robot-related, and environmental (Hancock et al., 2011; Hancock et al., 2020; Schaefer, 2013). Since this research analyzes the psychosocial factors – contemplating that an individual can be impacted by bottom-up (input from external reality) and top-down (input from cognitive, behavioral, and affective, therefore internal) processes – if a variable influences trust and since trust is inherently psychological, it will be considered a psychosocial factor. This research aims, consequently, to map and integrate in dimensions of categories every significant psychosocial variable.

This systematic literature review aims to gather and integrate relevant information concerning psychosocial variables that impact trust in HRI within the scope of the industrial context, and with special emphasis on collaborative robots or cobots. This literature review is, therefore, oriented by the following research question:

What are the psychosocial factors associated with trust in HRI in the industrial context?

HRI is often studied taking into consideration different categories of robotics depending on their usage – e.g., military robotics, industry robotics, medical robotics, high risk robotics, entertainment robotics, social robotics, and education robotics (Schaefer, 2013). This research considers only studies that can have clear implications on workers from industrial contexts. This narrowed the scope of the research and, consequently, the number of studies found.

This narrowed scope, along with the broad definition of concepts such as psychosocial factors and HRI revealed 1) a short number of articles matching the research criteria and 2) a fuzziness concerning the conceptual and pragmatic usability of several studies. Taking this into account, our study aimed to clarify concepts related to this topic and compile the research made from 2015 until the end of January 2021 to map the different psychosocial variables analyzed in the different studies. The accessibility of this kind of research can have a positive impact on organizational decision-making processes. Kopp and colleagues (2021) work, on this regard, underlines how understanding and acting in accordance with the impact trust can have on HRI can be a challenge for company representatives and decision-makers in the production sector.

Firstly, the main concepts will be defined to clarify the meaning of those terms here on. Thereafter the method used to do the search will be explained, mentioning the databases used and the quantitative metrics that result from the implementation of the research plan. After this, the research question will be answered using the articles found through the databases and doing the mapping of the psychosocial factors impacting trust that were considered. Done this part, the results and discussion will highlight the quantitative and qualitative results of this systematic literature review the limitations and implications of this study. Finally, the conclusion will summarize the main findings and present future research that can be done in this field.

2. Main Concepts

2.1. Human-Robot Interaction

Human-Robot Interaction (HRI) is defined by Fong et al. (2003) as "the study of the humans, robots, and the ways they influence each other" (p. 265). Therefore, this concept is based on the possible interactions in a Human-Robot Team (HRT). An HRT is constituted by at least one human – intelligent agent – and one robot - Artificial Intelligence (AI) or autonomous system (de Visser et al., 2020). Bauer and colleagues (2016) present five types of interaction that can be considered between human and robot and, therefore, considered typologies of HRI: 1) cell - robot is operating in a cage; 2) coexistence - the parts work together without sharing a workspace; 3) synchronized - they share the workspace in one at the time; 4) cooperation - they work in the same workspace at the same time, but not in the same product; 5) collaboration - both parts work in the same workspace, at the same time in the same product. HRI is relevant since it expands the work capacity and security of humans since automation, from which robotics is a subset, generates the possibility of executing operations only done before by humans or even impossible to them right now (Parasuraman et al., 2000).

2.2. Cobot

A cobot – or collaborative robot – is a robot that shares the worker's workspace and becomes their co-worker, having their own tasks which implies limited or no interaction – just coexistence - with the human to which they were assigned (Bergman et al., 2019). It is expected that this relationship evolves in a near future to cooperation (Bergman et al., 2019). A cobot can, therefore, be considered a robot created to collaborate or interact with humans in a shared workspace (Hentout et al., 2019).

2.3. Psychosocial Factors

Psychosocial factors, as mentioned by Suzuki and Takei (2013) are "influences that affect a person psychologically or socially" (p. 1582) and this concept includes mood status, cognitive behavioral responses, and social factors. These kinds of variables are related to a large range of examples linked to a person's psychological state and social environment, being conceptualized as processes and meanings at the individual level that can impact mental states (Upton, 2013). Factors such as locality, physical attributes, relationships with others and changes in personal roles should be taken into consideration (Suzuki & Takei, 2013). Therefore, we may consider for the purpose of this article, psychosocial factors as any variable that impacts psychologically the individual.

2.4. Trust

Trust is an attitude that, in uncertain and vulnerable situations, an agent will help in the achievement of one's personal goals (Lee & See, 2004). This attitude is crucial when concerning the development of relationships, not just amongst human beings, but also when they are based on human-technology interaction (Schaefer et al., 2016). Taking into consideration this interaction, trust can 1) help reduce risk, uncertainty, and anxiety, while 2) helping to create meaningful and positive experiences with technology, and 3) it is fundamental to make a user create and sustain a gradual interaction with the system (Gulati et al., 2019). Therefore, trust is relevant when operating a robotic system and can determine its acceptance and usage (Yagoda & Gillan, 2012) and designing trustworthy human-artificial intelligence (AI) interactions shows itself relevant to generate positive user experience (Lee et al., 2019) decreasing the probability of human rejection. This can be summarized in the principle "no trust, no use" (p. 377) presented by Schaefer et al. (2016) which refers to the fact that the rates of usage of automated systems by users are positively and proportionally linked to the trust that the user has on these automated systems.

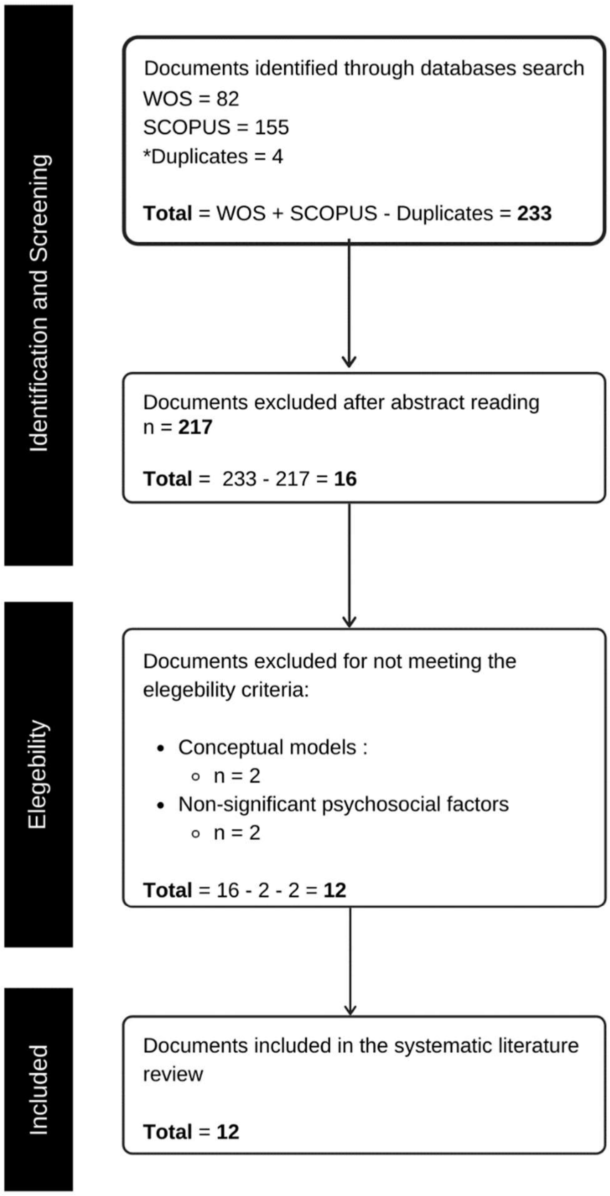
3. Method

As Denyer and Tranfield (2009) mention, a systematic review of literature should explore a clearly specified question. The search protocol follows the guidelines of the research plan presented by these authors. To find articles for the systematic literature review two databases were used: Web of Science and Scopus. On February 13th, 2021 we filtered the studies by articles, reviews, and proceedings papers from 2015 to 2021 published in English.

The query string used was the following: (cognitive OR mental OR emotion* OR psycho* OR behav*) AND (criteri* OR indicat* OR determinant* OR factor* OR variable OR predictor) AND (trust) AND (human AND robot AND interact* OR human-robot AND interact* OR human-interact* AND robot OR human-interact* OR hri OR collab* AND robot OR collab* AND human-robot OR human AND robot AND collab* OR human-robot AND collab* AND workstation* OR co-robotic OR cobot OR cobotic OR light AND robot* OR human AND robot AND cooperat* OR human-robot AND cooperat* OR human-cooperat* AND robot OR hci) AND (manufactur* OR factor* OR producti* OR industr* OR lab*). This contemplates the keywords found in the research question and synonyms or keywords strongly related to them. Concerning specifically the keyword "psychosocial" it was considered that behavior, psychology, cognition, mental and emotion were relevant concepts to try to fit in this string, since they can add information related to studies which fit within our scope of research. As eligibility criteria conceptual models were excluded and only significant variables were included. From the eighty-two results on Web of Science database, nine were valid for the purpose of the research. From the 155 results present on Scopus, three were considered, being one of them also in Web of Science. Summarizing, from 233 different articles found, a total of 12 articles were considered valid for the purpose of the systematic literature review.

Figure 1

Flowchart representing the selection process of the final sample



4. Results

Using the definition of Suzuki and Takei (2013) this systematic literature review is going to map the psychosocial factors in dimensions of 1) mood status 2) cognitive behavioral responses, 3) social factors, and/or 4) others. Hereupon, the studies considered in methodology – 12 articles – will be analyzed to understand which of these dimensions of factors they fit into. Taking this into consideration makes it possible to map the factors considered on Web of Science and Scopus from 2015 to 2021 and understand 1) if they match these dimensions and 2) if they include new variables that were not contemplated on them. It is also relevant to understand what variables were significant.

Fratczak and colleagues' (2019) proceeding – simulation – reveals that, not having impact on performance of the human, the robot's behavior influences his/her trust and shows that trust is regained naturally over time but the rate of progression for this to happen depends on how the robot behaves. This experiment was made using means of Virtual Reality (VR) which classifies it as a simulation. Robot's behavior shall be categorized belonging to both the cognitive behavioral responses and social factors dimensions, since it seems to fit the idea of a response – we do not have enough information to define if it is cognitive or behavioral – to an external stimulus and this stimulus is created by the robot's behavior during its relationship with her/his user. Time shall be added to the category of "others" since it is a variable that is not directly social or psychological.

In Sadrfaridpour and Wang (2018) simulation, physical HRI (pHRI) and social HRI (sHRI) were analyzed. The proceeding concluded that both pHRI and sHRI impact trust together, but pHRI alone does not do so. The clarification of the robot's intent and predictability improves trust. Both pHRI and sHRI shall be added to the social factors dimension since they are firstly focused on categories of HRI.

In their article and by doing a simulation, Washburn and colleagues (2020) refer that Functionality framing – the perception on how the robot is functioning – impacts trust if it is low when the robot makes errors and does not impact trust if, in the same situation, that variable is high. This factor is considered a cognitive behavioral response since it is a perception of how the robot works – being congruent with the attribution of expectancy done before – but it is relevant to disclaim that this framing construction can be influenced by other variables such as social factors.

In another article applying a simulation, Macarthur and colleagues (2017) took into consideration proximity and speed as variables and both shown to be significant and, therefore, impact trust. Both these variables will be considered as social factors, because even though they impact trust through individual responses, they are ultimately factors related to physical location and movements of the robot.

Agreeable trait has a significant influence on the formation of trust in HRI (Huang et al., 2020). The variable found in this proceeding and experimented through a simulation shall be added to the dimension mood status since it is an individual trait present in the individual's personality and not a response or a social factor.

Natarajan and Gombolay' experiment in a simulated environment (2020) tested the following factors and their impact on trust: perceived anthropomorphism (significant), robot presence (non-significant), robot type (non-significant), robot behavior (significant). Considering only the ones on this proceeding that are significant to impact trust, it is possible to state that the perceived anthropomorphism is both a cognitive behavioral response and a social factor – it has to do with the physical attribute of the robot but here it is considered the way it is perceived by the user – and robot behavior is also both a cognitive behavioral response and a social factor since the authors found significance between behaviors of the robot when impacting the trust of the user.

Lohani and colleagues (2016) mention on their proceeding the relevance of mentalizing propensity – perception of sensors' role as facilitators to mentalization (this process allows a human to make sense of his/her mental states, as well as of the mental states of others) – to impact positively trust and behavioral reliance. Therefore, this simulation reveals social interactions with a robot impacts the user's mentalizing propensity and this suggests that a robot's social behavior can influence the perceived role of sensors used in HRI impacting consequently the human-robot trust. This can classify mentalizing propensity as both a cognitive behavioral response and social factors simultaneously, since the top-down (cognitive attribution of social roles and behaviors to a robot) and bottom-up processes (the congruence of the behavior of the robot with this perception) generate this variable.

A simulation made by Palmarini and colleagues (2018) concludes that the context-awareness or the workspace awareness is relevant to provide the confidence needed to enhance the perception of safety, impacting positively her/his trust in his/her interaction with the robot. In this sense, this article shows that human safety and context awareness are psychosocial factors that impact trust and shall be added to both the cognitive behavioral response and the social factors dimensions, since they are linked with perception of safety and the attributes of the environment/robot.

Wang and colleagues (2019) mention a weighting factor (λ) that incorporates the intention of a user to interact with a specific robot. This variable was found through a simulation to have a significantly negative correlation with trust being possible to explain the later concept with the intention of interaction with a robot – but the significance was just found in long-range tasks. This weighting factor referred in this article shall be considered both a cognitive behavioral response – due to the intention component of it – and a social factor – due to its interaction component.

Agarwal and Bharti (2019) refer in their article done through engaging engineers and final users to discuss and answer questionnaires about autonomous mobile robots (AMR) that reliability is a relevant factor impacting trust in technology, but relative advantage, self-efficiency and performance expectancy are also important. Reliability can be classified as both a social factor and a cognitive behavioral response, relative advantage as a social factor, and both self-efficiency and performance expectancy as cognitive behavioral responses.

According to the Meta-Analyses performed by Hancock and colleagues (2020), satisfaction, expectancy, comfort with robots, level of complacency, tendency to anthropomorphize all objects, level of extroversion, reliability and robot personality were found to have a significant impact on trust, using correlational analyses. Factors such as human operator's prior experience, attitude to robots, propensity to trust, race, age, gender, expertise, self-efficacy, education, culture, dependability, appearance, anthropomorphism, behavior, communication, in-group membership, task complexity, and interaction frequency were not significant in this context. However, according to the conducted pairwise comparisons, human operator's prior experience, age, culture, and anthropomorphism significantly impacted trust along with other factors such as task difficulty, robot reliability, proximity, and experience/tenure. Thus, the degree of trust was higher when tasks were more difficult, when the robot was more reliable, when the location was closer, when there was more experience and higher degrees of anthropomorphism, when the participants were younger, and for Asian cultures and the United States in comparison to European cultures. Factors such as robot behavior, level of automation (LOA), feedback, and task type were not significant, in accordance with the pairwise comparisons. In respect to collaborative tasking, the article showed a tendency for users to trust higher a robot that belonged to their group than one that did not. All these factors were divided into subsections of three categories. According to the correlational analyses, the overall human-related antecedents and robot-related antecedents were significantly related to trust in a robot, while the overall contextual antecedents were not. The characteristic subsection of the human-related antecedents was significant, but, however, the ability-based one was not. Both subsections for robot-related antecedents, namely, performance-based features and attribute-based features were significant. Finally, team collaboration, included in the category of contextual antecedents, was significant, while team's current task was not. However, the results for the pairwise comparisons were different, indicating that all the overall categories had a significant impact on trust, and that, of the human-related antecedents, human characteristics, but not ability-based factors, were significant, of the robot-related antecedents, both subcategories of factors proved to be significant and, of the contextual antecedents, none of the subcategories was significant. The authors noted that the impact on trust was stronger for factors relating to the robot in comparison to factors relating to the human. Dimensioning all the significant factors on this study, satisfaction, expectancy and level of complacency shall be added to cognitive behavioral responses, level of extroversion shall be added to mood status, culture, robot personality, anthropomorphism, task difficulty and proximity shall be added to social factors, prior experience, comfort with robots, tendency to anthropomorphize objects and reliability shall be considered both cognitive behavioral responses and social factors and, finally, age shall fit in “others” category since it is a temporal construct that isolated is not relatable with the other categories.

In a literature review developed by Honig and Oron-Gilad (2018) several variables were identified. The first of them was related to errors and the context they occur. The authors noticed that the negative impact on trust was higher when the errors were done by a conversational speech robot in comparison to one that was just functional (Cha et al., 2015, cited in Honig & Oron-Gilad, 2018). The same linkage between errors and trust can be observed when comparing visual aesthetic user interface with non-visual aesthetic user interface: the first one has a more negative impact on trust compared to the second one (Weinstock et al., 2012, cited in Honig & Oron-Gilad 2018). If the consequences of the error are severe, the loss of trust is greater than it would be if they were not (Rossi et al., 2017, cited in Honig & Oron-Gilad 2018). Another variable is reliability and concerning it, Desai and colleagues (2013, cited in Honig & Oron-Gilad, 2018) stated in their literature review that when low reliability starts sooner for a subject there is a higher negative impact of this variable on trust comparing to one having low reliability in later periods. However, the same authors reached a contrasting finding in a preliminary study one year before (Desai et al., 2012, cited in Honig & Oron-Gilad, 2018) where they mentioned that less trust in the robot was showed when the level of reliability of a subject on the object drops later in time or in the middle run. Honig and Oron-Gilad (2018) also mentioned that timing, consequence, and expectations are factors associated with failure that impact trust. Also, if the robot apologizes, expresses regret or reparation, the trust ratings are similar as a non-failing robot (Hamacher et al., 2016, cited in Honig & Oron-Gilad 2018). If the robot places blame for a failure it reduces the user’s trust (Kaniarasu & Steinfeld, 2014, cited in Honig & Oron-Gilad 2018). In this same logic, demonstrating appropriate emotions and awareness of error makes the agent more trustful in the eyes of the user (Hamacher et al., 2016, cited in Honig & Oron-Gilad 2018). Errors were categorized as both cognitive behavioral responses and social factors, since this concept, when mentioned in this literature review, is a cognitive response to a perceived role of a machine. Reliability shall also be added to both cognitive behavioral response and social factors, since it is linked to the cumulative response given by an individual to an attribute of the robot in a temporal period. Timing and consequence are social factors that influence the user, as they depend, respectively, on the temporal frame and impact of the action of an external object. Expectation is firstly a cognitive behavioral response. The robot apologizing, expressing regret or reparation, placing blame, and showing appropriate emotions shall be considered both a social factor and cognitive behavioral responses since it is generated by the action of the object and the role given to it by the human – anthropomorphizing the object – which, simultaneously, will trigger empathy for the robot.

5. Discussion

From the 12 articles included, one was based on the results of discussions and questionnaires done to engineers and potential users of industries. Therefore, this study was considered done in an industrial environment. Other 9 studies were considered done in simulated environments – there were found presential simulations and Virtual Reality (VR) simulations. We did not apply this analysis to the two reviews. Five documents were proceedings, five articles and two of them were reviews. Concerning the publishing year, from 2015 to early 2021, two documents are from 2016, three from 2018, three from 2019 and four from 2020. The low results for 2021 can be caused by the fact that this systematic review is done in an early period of this year (see Table 1).

Table 1

Distribution of Documents by Type, Research Environment, and Publishing Year

	Type			Environment				
Documents	Article	Proceeding	Review	Industrial	Simulate	N.a.	Year	Quantity
(Hancock et al.,2020)			x			x	2020	4
(Huang et al.,2020)		x			x			
(Natarajan & Gombolay,2020)		x			x			
(Washburn et al.,2020)	x				x			
(Agarwal & Bharti, 2019)	x						2019	3
(Fratczak et al., 2019)		x			x			
(Wang et al., 2019)	x			x	x			
(Honig & Oron-Gilad,2018)			x			x	2018	3
(Palmarini et al., 2018)	x				x			
(Sadrfaridpour & Wang,2018)		x			x			
(Lohani et al.,2016)		x			x		2016	2
(Macarthur et al.,2016)	x				x			

From the studies analyzed a total of 36 different psychosocial variables were found (see Table 2). Taking into consideration the dimensions defined initially from the variables mentioned – and accepting that each one can be inserted in any adequate number of dimensions – two were added to mood status (MS), 23 added to cognitive behavioral responses (CBR), 25 to social factors (SF) and two to “others”. This mapping of psychosocial factors inside dimensions shows a quantitative impact of social factors and behavioral responses on trust concerning HRI. Several variables – 16 out of 36 – were both classified as cognitive behavioral responses and social factors which reveal the necessity of considering a psychosocial approach when studying the effects of these variables on trust.

Table 2

Categorization of Psychosocial Factors in Dimensions

		Dimensions				
Documents	Psychosocial Factors	MS	CBR	SF	Others	Number of repetitions
(Natarajan & Gombolay, 2020)	Perceived anthropomorphism		X	X		x
	Robot behavior		X	X		xx
(Fratczak et al., 2019)	Robot behavior		X	X		xx
	Time				X	x
(Sadrfaridpour & Wang, 2018)	pHRI			X		x
	sHRI			X		x
(Washburn et al., 2020)	Functionality framing		X			x
(Macarthur et al., 2016)	Proximity			X		xx
	Speed			X		x
(Huang et al., 2020)	Agreeable trait	X				x
(Honig & Oron-Gilad, 2018)	Errors		X	X		x
	Reliability		X	X		xxx
	Timing		X	X		x
	Consequence		X	X		x
	Expectations		X			x
	Apologizing, regretting or reparation		X	X		x

		Dimensions				
Documents	Psychosocial Factors	MS	CBR	SF	Others	Number of repetitions
	Place blame		X	X		x
	Appropriated emotions		X	X		x
(Lohani et al., 2016)	Mentalizing propensity		X	X		x
(Palmarini et al., 2018)	Human safety		X	X		x
	Context awareness		X	X		x
(Wang et al., 2019)	Weighting factor		X	X		x
(Agarwal & Bharti, 2019)	Reliability		X	X		xxx
	Relative advantage			X		x
	Self-efficiency		X			x
	Performance expectancy		X			x
(Hancock et al., 2020)	Prior experience		X	X		x
	Satisfaction		X			x
	Expectancy		X			x
	Comfort with robots		X	X		x
	Level of complacency		X			x
	Tendency to anthropomorphize objects		X	X		x
	Level of extroversion	X				x
	Age				X	x
	Culture			X		x

		Dimensions				
Documents	Psychosocial Factors	MS	CBR	SF	Others	Number of repetitions
	Reliability		X	X		xxx
	Robot personality			X		x
	Anthropomorphism			X		x
	Task difficulty			X		x
	Proximity			X		xx
Total (Without Duplicates)		2	23	25	2	
Documents	Different psychosocial factors					
12	36					

Note. This table contemplates the distribution of the psychosocial factors over the established dimensions, namely, mood status (MS), cognitive behavioral responses (CBR), social factors (SF) and "others". Since some of the factors are duplicated along the table, the number of "x" in the last column indicates the number of times the variables appear throughout the table.

5.1. Limitations

The exclusion of several variables due to not being proven to be significant, reduced greatly the number of studies and, consequently, the number of psychosocial factors. This limitation was a tradeoff to augment the certainty of data that was included in this study and to enable the attribution of significant variables to dimensions. Another possible limitation refers to the inclusion of psychosocial factors in dimensions, evaluating them conceptually by our understanding in the absence of a developed response from the original authors about their conceptualization of variables. This generates possible interpretation gaps between what could be the conceptual reasoning of the authors and ours concerning the psychosocial factors on the table. It is relevant to consider that this map shall be updated if more epistemological information is added to the variables. The studies considered were mainly simulations which may present another limitation since the participants were not experimenting in the industrial context and some of them were not even industrial workers. This can create a bias in certain results and may need a confirmatory experiment done in industries.

5.2. Implications

Despite what was mentioned before, this study was able to map several factors impacting trust and, consequently, can be used as a tool to help future mapping processes related to this field. This can facilitate the understanding of the psychosocial subcategory of the new variables and act on them through the implementation of strategies, development of studies and experiments, knowing more about what type of processes are being contemplated and how to deal with better results considering their nature and dynamics. Understanding the source of the input can generate a better comprehension of its influence in the generation of a certain output. Another relevant point of this study is the definition of main concepts that are vague or guaranteed as self-explanatory by several other studies. The conceptual ambiguity presented in concepts like psychosocial factors and human-robot interaction can have a negative impact on the possibility to create consensus to map factors that need to be dimensioned to facilitate the design of robotics considering the reaction of the user. The mapping of psychosocial variables taking into account the impact of internal and external processes and their combination can guide future studies to further understand why these factors impact trust, enabling the implementation of projects more appropriate to this subject and improving performance, safety and well-being of workers in Industry 4.0.

6. Conclusion

Several psychosocial factors that have impact on trust in HRI were detected and mapped in this study. From these factors only two were considered related to mood status and two classified as “others”. From these two we can see a pattern of linkage to a temporal dimension, which can justify the creation of a new dimension for this kind of variable. Approximately 44,44% (16 out of 36) of all the variables were considered cognitive behavioral responses and social factors at the same time, showing the need of a psychosocial approach to study and correctly implement variables with influence on trust. The temporal range for this systematic literature review can limit the possibility of having more results, but since this study aims to consider the impact of robotics on trust in industrial context, constant technological updates within Industry 4.0 shall be considered and the possible risk of contemplating studies based on variables such as outdated technology should be avoided. This model has just four categories and future research can enhance its capacity to map this reality. One possible suggestion is creating subdimensions such as temporal factors or even demographic factors to understand more interdependencies and specifications, creating conditions for deepening knowledge and detecting similarities. Also, being automation and artificial intelligence, larger categories where robotics can be inserted, it would be relevant to understand what in Human-Computer Interaction and Human-Automation Interaction can be applied to HRI. It would also be important to understand if factors that impact other categories of robots – e.g., social robots – can be extrapolated to robots in the industrial context.

References

Agarwal, D., & Bharti, P. S. (2019). Computation of cause and effect relationship for acceptance of autonomous mobile robots in industries. *Journal of Statistics and Management Systems*, *22*(2), 237–256. https://doi.org/10.1080/09720510.2019.1580903.

Ahuett-Garza, H., & Kurfess, T. (2018). A brief discussion on the trends of habilitating technologies for Industry 4.0 and Smart manufacturing. *Manufacturing Letters*, *15*, 60–63. https://doi.org/10.1016/j.mfglet.2018.02.011.

Bauer, W., Bender, M., Braun, M., Rally, P., & Scholtz, O. (2016). Lightweight robots in manual assembly-best to start simply. Stuttgart: Frauenhofer-Institut für Arbeitswirtschaft und Organisation IAO.

Bednar, P. M., & Welch, C. (2020). China1. *Information Systems Frontiers*, *22*(2), 281–298.

de Visser, E. J., Peeters, M. M. M., Jung, M. F., Kohn, S., Shaw, T. H., Pak, R., & Neerincx, M. A. (2020). Towards a Theory of Longitudinal Trust Calibration in Human–Robot Teams. *International Journal of Social Robotics*, *12*(2), 459–478. https://doi.org/10.1007/s12369-019-00596-x.

Denyer, D., & Tranfield, D. (2009). Producing a systematic review. In D. A. Buchanan & A. Bryman (Eds.), The SAGE handbook of organizational research methods. In *The Sage handbook of organizational research methods* (pp. 671–689). https://www.mendeley.com/catalogue/0ff0533c-01f2-338e-a1a9-6f60a28fa801/?utm_source=desktop&utm_medium=1.19.4&utm_campaign=open_catalog&userDocumentId=%7B8b821bd1-4e6e-4cce-8efd-130b29a0bf5b%7D.

Fong, T., Thorpe, C., & Baur, C. (2003). Collaboration, Dialogue, Human-Robot Interaction. In R. A. Jarvis. & A. Zelinsky (Eds.), R*obotics Research. Springer Tracts in Advanced Robotics* (vol 6, pp. 255-266). Springer, Berlin, Heidelberg. https://doi.org/10.1007/3-540-36460-9_17.

Fox, W. M. (1995). Sociotechnical System Principles and Guidelines: Past and Present. *The Journal of Applied Behavioral Science*, *31*(1), 91-105. https://doi.org/10.1177/0021886395311009.

Fratczak, P., Goh, Y. M., Kinnell, P., Soltoggio, A., & Justham, L. (2019). Understanding Human Behaviour in Industrial Human-Robot Interaction by Means of Virtual Reality. *HTTF 2019: Proceedings of the Halfway to the Future Symposium 2019*. 1-7. https://doi.org/10.1145/3363384.3363403.

Gulati, S., Sousa, S., & Lamas, D. (2019). Design, development and evaluation of a human-computer trust scale. *Behaviour and Information Technology*, *38*(10), 1004–1015. https://doi.org/10.1080/0144929X.2019.1656779.

Hancock, P. A., Kessler, T. T., Kaplan, A. D., Brill, J. C., & Szalma, J. L. (2020). Evolving Trust in Robots: Specification Through Sequential and Comparative Meta-Analyses. *Human Factors*. https://doi.org/10.1177/0018720820922080.

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

Hentout, A., Aouache, M., Maoudj, A., & Akli, I. (2019). Human–robot interaction in industrial collaborative robotics: a literature review of the decade 2008–2017. *Advanced Robotics*, *33*(15–16), 764–799. https://doi.org/10.1080/01691864.2019.1636714.

Honig, S., & Oron-Gilad, T. (2018). Understanding and resolving failures in human-robot interaction: Literature review and model development. *Frontiers in Psychology*, *9*(JUN). https://doi.org/10.3389/fpsyg.2018.00861.

Huang, Hsiao-Ying & Twidale, Michael & Bashir, Masooda. (2020). 'If You Agree with Me, Do I Trust You?': An Examination of Human-Agent Trust from a Psychological Perspective. In Y. Bi, R. Bhatia, & S. Kapoor (Eds.) *Advances in Intelligent Systems and Computing* (vol. 1038). Springer, Cham. https://doi.org/10.1007/978-3-030-29513-4_7310.1007/978-3-030-29513-4_73.

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

Lee, Y., Ha, M., Kwon, S., Shim, Y., & Kim, J. (2019). Egoistic and altruistic motivation: How to induce users' willingness to help for imperfect AI. *Computers in Human Behavior*, *101*(July), 180–196. https://doi.org/10.1016/j.chb.2019.06.009.

Lohani, M., Stokes, C., McCoy, M., Bailey, C. A., Joshi, A., & Rivers, S. E. (2016). Perceived role of physiological sensors impacts trust and reliance on robots. *25th IEEE International Symposium on Robot and Human Interactive Communication, RO-MAN 2016*, 513–518. https://doi.org/10.1109/ROMAN.2016.7745166.

Macarthur, K. R., Stowers, K., & Hancock, P. A. (2017). Human-robot interaction: Proximity and speed—slowly back away from the robot! *Advances in Intelligent Systems and Computing, 499*, 365–374. https://doi.org/10.1007/978-3-319-41959-6_30.

Natarajan, M., & Gombolay, M. (2020). Effects of anthropomorphism and accountability on trust in human robot interaction. *ACM/IEEE International Conference on Human-Robot Interaction*, 33–42. https://doi.org/10.1145/3319502.3374839.

Palmarini, R., Del Amo, I. F., Bertolino, G., Dini, G., Erkoyuncu, J. A., Roy, R., & Farnsworth, M. (2018). Designing an AR interface to improve trust in Human-Robots collaboration. *Procedia CIRP*, *70*, 350–355. https://doi.org/10.1016/j.procir.2018.01.009.

Parasuraman, R., Sheridan, T. B., & Wickens, C. D. (2000). A model for types and levels of human interaction with automation. *IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans*, *30*(3), 286-297. https://doi.org/10.1109/3468.844354.

Pasmore, W., Winby, S., Mohrman, S. A., & Vanasse, R. (2019). Reflections: Sociotechnical Systems Design and Organization Change. *Journal of Change Management*, *19*(2), 67–85. https://doi.org/10.1080/14697017.2018.1553761.

Sadrfaridpour, B., & Wang, Y. (2018). Collaborative Assembly in Hybrid Manufacturing Cells: An Integrated Framework for Human-Robot Interaction. *IEEE Transactions on Automation Science and Engineering*, *15*(3), 1178–1192. https://doi.org/10.1109/TASE.2017.2748386.

Schaefer, K. (2013). The Perception and Measurement of Human-Robot Trust. *University of Central Florida (Doctoral Dissertation)*, 360.

Schaefer, K. E., Chen, J. Y. C., Szalma, J. L., & Hancock, P. A. (2016). A Meta-Analysis of Factors Influencing the Development of Trust in Automation: Implications for Understanding Autonomy in Future Systems. *Human Factors*, *58*(3), 377–400. https://doi.org/10.1177/0018720816634228.

Sung, T. K. (2018). Industry 4.0: A Korea perspective. *Technological Forecasting and Social Change*, *132*(October 2017), 40–45. https://doi.org/10.1016/j.techfore.2017.11.005.

Suzuki, S., & Takei, Y. (2013). Psychosocial Factors and Traumatic Events. In M. D. Gellman & J. R. Turner, *Encyclopedia of Behavioral Medicine* (pp. 1582-1583). https://doi.org/10.1007/978-1-4419-1005-9.

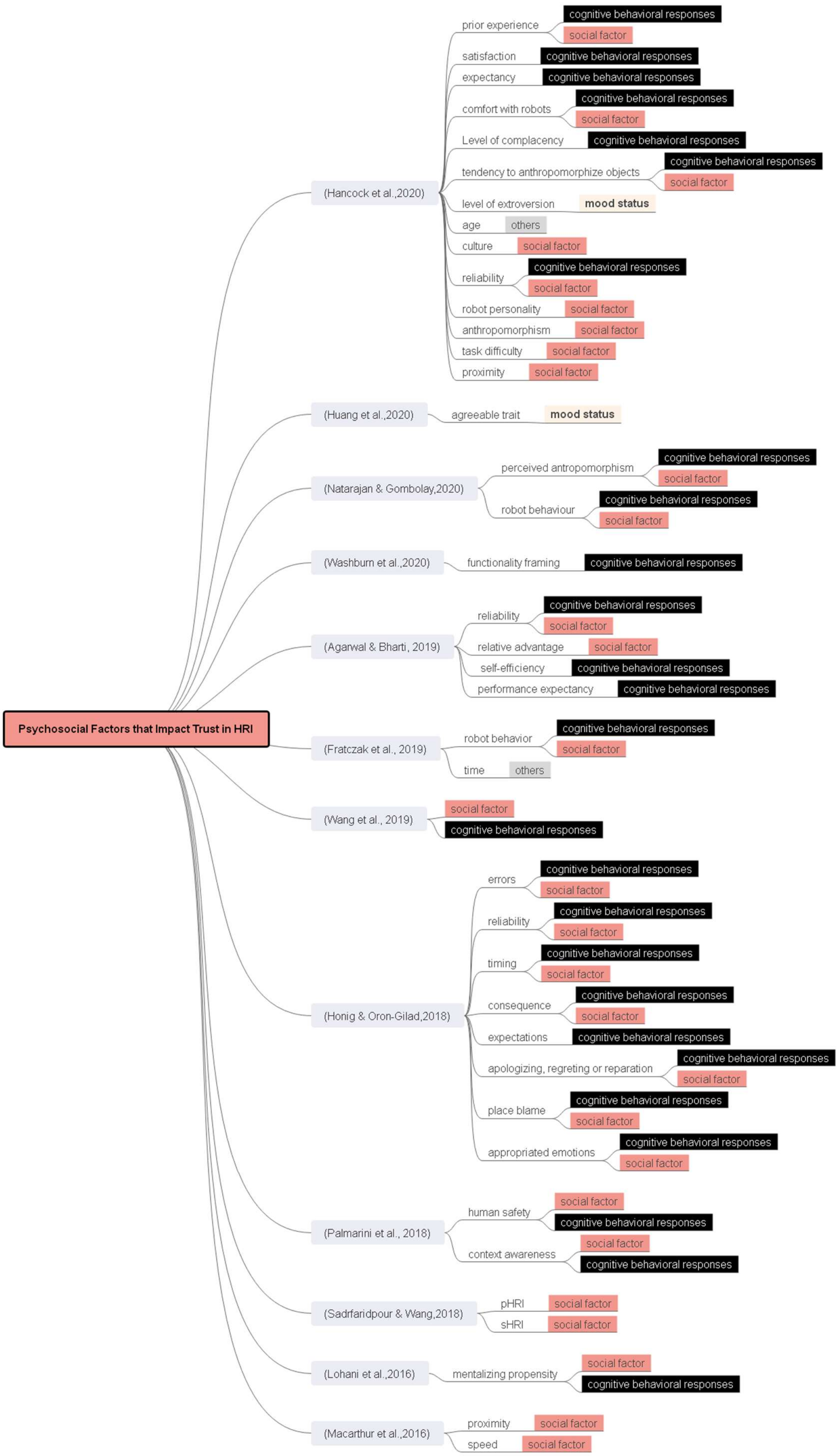
Upton, J. (2013). Psychosocial Factors. In M. D. Gellman & J. R. Turner, *Encyclopedia of Behavioral Medicine* (pp. 1580-1581). https://doi.org/10.1007/978-1-4419-1005-9.

Wang, Y., Lematta, G. J., Hsiung, C.-P., Rahm, K. A., Chiou, E. K., & Zhang, W. (2019). Quantitative Modeling and Analysis of Reliance in Physical Human–Machine Coordination. *Journal of Mechanisms and Robotics*, *11*(6). https://doi.org/10.1115/1.4044545.

Washburn, A., Adeleye, A., An, T., & Riek, L. D. (2020). Robot errors in proximate hri how functionality framing affects perceived reliability and trust. *ACM Transactions on Human-Robot Interaction*, *9*(3). https://doi.org/10.1145/3380783.

Yagoda, R. E., & Gillan, D. J. (2012). You Want Me to Trust a ROBOT? The Development of a Human-Robot Interaction Trust Scale. *International Journal of Social Robotics*, *4*(3), 235–248. https://doi.org/10.1007/s12369-012-0144-0.

APPENDIX


Psychosocial Factors that Impact Trust in HRI
(Hancock et al.,2020)
prior experience
cognitive behavioral responses
social factor
satisfaction
cognitive behavioral responses
expectancy
cognitive behavioral responses
comfort with robots
cognitive behavioral responses
social factor
Level of complacency
cognitive behavioral responses
tendency to anthropomorphize objects
cognitive behavioral responses
social factor
level of extroversion
mood status
age
others
culture
social factor
reliability
cognitive behavioral responses
social factor
robot personality
social factor
anthropomorphism
social factor
task difficulty
social factor
proximity
social factor
(Huang et al.,2020)
agreeable trait
mood status
(Natarajan & Gombolay,2020)
perceived antropomorphism
cognitive behavioral responses
social factor
robot behaviour
cognitive behavioral responses
social factor
(Washburn et al.,2020)
functionality framing
cognitive behavioral responses
(Agarwal & Bharti, 2019)
reliability
cognitive behavioral responses
social factor
relative advantage
social factor
self-efficiency
cognitive behavioral responses
performance expectancy
cognitive behavioral responses
(Fratczak et al., 2019)
robot behavior
cognitive behavioral responses
social factor
time
others
(Wang et al., 2019)
social factor
cognitive behavioral responses
(Honig & Oron-Gilad,2018)
errors
cognitive behavioral responses
social factor
reliability
cognitive behavioral responses
social factor
timing
cognitive behavioral responses
social factor
consequence
cognitive behavioral responses
social factor
expectations
cognitive behavioral responses
apologizing, regreting or reparation
cognitive behavioral responses
social factor
place blame
cognitive behavioral responses
social factor
appropriated emotions
cognitive behavioral responses
social factor
(Palmarini et al., 2018)
human safety
social factor
cognitive behavioral responses
context awareness
social factor
cognitive behavioral responses
(Sadrfaridpour & Wang,2018)
pHRI
social factor
sHRI
social factor
(Lohani et al.,2016)
mentalizing propensity
social factor
cognitive behavioral responses
(Macarthur et al.,2016)
proximity
social factor
speed
social factor