



## Open Research Online

### Citation

Huang, Xinyu and Wild, Fridolin (2021). A new metric scale for measuring trust towards holographic intelligent agents. In: Proceedings of the 34th British HCI Workshop and Doctoral Consortium 2021, BCS Learning and Development Ltd.

### URL

<https://oro.open.ac.uk/82555/>

### License

(CC-BY 4.0) Creative Commons: Attribution 4.0

<https://creativecommons.org/licenses/by/4.0/>

### Policy

This document has been downloaded from Open Research Online, The Open University's repository of research publications. This version is being made available in accordance with Open Research Online policies available from [Open Research Online \(ORO\) Policies](#)

### Versions

If this document is identified as the Author Accepted Manuscript it is the version after peer review but before type setting, copy editing or publisher branding

# A new metric scale for measuring trust towards holographic intelligent agents

Xinyu Huang and Fridolin Wild  
Institute of Educational Technology,  
The Open University, UK  
{xinyu.huang,f.wild}@open.ac.uk

**Trust is an essential attitude in social relationships, but it also mediates our approach to certain technology. The definition of interpersonal trust, however, is too wide to expound our understanding of how trust impedes such interaction with technology, and the lack of an applicable quantifiable model in particular presents an obstacle to our quest of building reliable, trusted, and intelligent holographic agents. In this paper, we therefore develop a novel metric scale to measure trust. We identify, select, and refine over a hundred items related to trust, check their precision and validity with the help of a judges panel, and select polarising items that are able to bring out the distinctive characteristics regarding people's trust towards intelligent agents. Our findings indicate that an assessment of trust involves looking at the user's belief about the agent's competence, integrity, benevolence, and compassion, which drive the attitude of trust, influenced by the user's general propensity to trust. Trust then drives intention to engage and ultimately engagement, which, if successful, results in the establishment of a trust relationship with the agent. While we propose an item-response scale for measuring this model of trust, we also add our deliberations on how elements of it could be replaced with alternative means that possibly offer more immediacy than self-inspection, discussing in particular alternatives for measuring elements of compassion, competence, and social relationships.**

*Human-Computer Interaction, Trust, Augmented Reality, Holographic Intelligence Agent, Holographic AIs.*

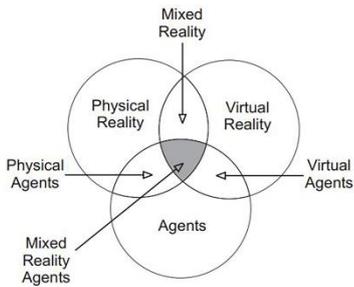
## 1. INTRODUCTION

With rise of compatible portable consumer devices (both mobile as well as head-mounted), Mixed Reality (MR) and Augmented Reality (AR) have become very commonplace in Human-Computer Interaction, not least due to their ability to integrate both digital and physical information. Inspired by their success in VR games, the use of agents in AR and MR is well investigated. Holz et al. (2011) propose a taxonomy for Mixed Reality Agents focused on agency, corporeal presence, and their interactive capacity (see Figure 1). They define a Mixed Reality agent (MiRA) as “an agent embodied in a mixed reality environment”, emphasising ‘embodiment’ in their circular definition and including robots. Based on that, Campbell et al. (2013) explore Augmented Reality agents (AuRAs) and their environmental contexts. Both definitions bring out how agents integrate into the physical world and what degree of behavioural realism they express. Although MiRAs emphasize the relationship of intelligence and adaptivity, the taxonomy does not indicate whether it is a real-time response nor further define intelligent features of virtual agents in the mixed context. In our view, spatial mapping and interactive/environmental responsiveness is key to distinguish holographic intelligent agents from MiRAs and AuRAs.

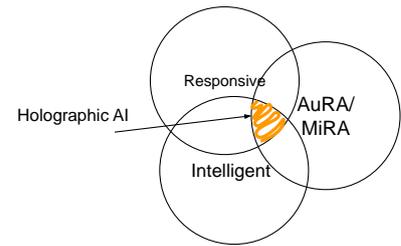
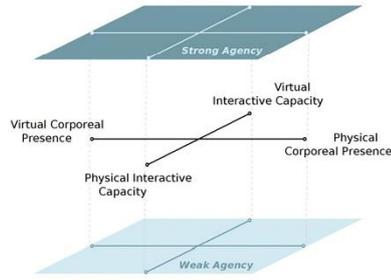
Besides, intelligence of virtual agents is more than just adapting with the help of embodiment to a

physical world, it also includes multimodal interaction (e.g., using natural language processing and dialogue understanding), as well as capability to perceive and match users' demands. Moreover, responsiveness requires the agent to be able to interact with both virtual and physical surroundings in real-time. In our view, intelligent holographic agents ('holographic AIs') should have capacity to react in real-time, driven by their analytical and inferential abilities and aimed at supporting the execution of the users' tasks in the mixed reality environment. Figure 2 represents this difference in a new Venn diagram.

With the help of Augmented Reality, we can project intelligent agents 'holographically' into a user's view of the physical surrounding. For such agents, we have coined the notion of 'Holographic Artificial Intelligences (AIs)'. They have humanoid appearance and behaviour, possess enough intelligence to adapt experiences to context, offering responsive face-to-face and environmental interaction (Huang, Wild, and Whitelock, 2021). Virtual agents are sometimes even treated as equivalent to interactive partners (Pinxteren et al., 2020). Positive interaction between users and holographic AIs, however, relies on their affective attachment and a rational assessment on whether to invest into a reciprocal relationship. Trust is a key driver, influencing this relationship.



**Figure 1.** A taxonomy of Mixed Reality agents (Holz et al., 2011).

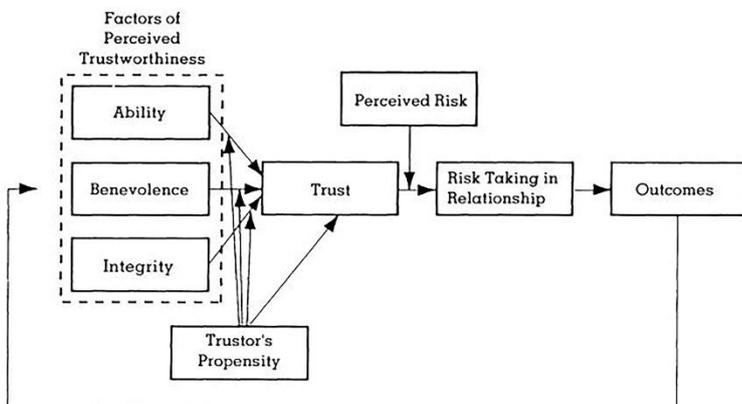


**Figure 2.** Holographic AI compared to AuRA/MiRA.

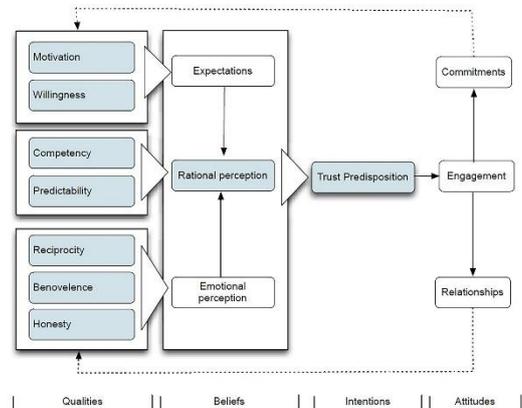
Mayer et al. (1995) propose competence, integrity, and benevolence as the essential elements to establish a model of trust. Computer agents are social actors (Nass et al. 1994). In human-computer interaction (HCI), competence means the computer agents possess necessary skills, abilities, and knowledge to execute and complete tasks. The definition of integrity is consistency of behaviour and promise in human-computer trust, and it allows systems to take responsibility and fulfil promises (Dobel, 1999). The process of evaluation towards integrity is rational by cognitive judgement. Benevolence reflects whether agents care about users, and execute tasks based on their interests (Phillip et al., 2020). Competence and integrity are objective evaluation criteria, and benevolence has an emotional component.

Cognitive perception is based on judgements or evaluations that make one feel confident in the other party's knowledge and ability (Borum, 2010). The cognitive aspect involves the trustors' belief that users believe agents are trustworthy in digital environments. Thus, behaviours and competence drive cognitive trust. The affective foundation, on the other hand, is emotion. Affective trust emerges from both the level of concern, caring, and benevolence (Borum, 2010).

Figure 3, from Mayer et al. (1995), depicts how perception of values influences the nature of trust. Values, attitudes, and emotions are three trust traits in interpersonal trust (Jones and George, 1998). In terms of HCI, values employ to shape agents' features or personalization to speculate user experience, and emotions influence attitudes and relationship. Trust involves intentions, perceptions, and beliefs. It relies on belief, and belief, behaviour, intention, and attitude also affect the degree of trust (Kulms and Kopp, 2018). Besides, trust reflects a state of mind, a willingness and predisposition, formed by an association of cognitive and affective based behaviours (Sousa, Lamas, and Dias, 2014, see Figure 4). Attitudes are emotional judgements of belief that are influenced by the accuracy of information and personal experiences in order to adopt intentions and produces corresponding behaviours (Lee et al., 2013). Similarly, trust in HCI is able to be an attitude that intelligent agents are willing to help achieve intentions as well (Kulms and Kopp, 2018). Although trust towards automation can be defined an attitude that computer agents help achieve goals in uncertain context (Lee and See, 2004), there is a little evidence to prove this definition can be used for humanoid holographic AIs. With this research project, we intend to investigate exactly this: what is trust towards a holographic AI and how can we



**Figure 3.** Integrative model of trust (Mayer et al., 1995).



**Figure 4.** A model of Human-Computer Trust (Sousa et al., 2014).

measure it using a questionnaire to quantify the sense of trust.

The work we present here focuses on identifying the relevant constructs that predict trust towards a holographic AI. For simplicity reasons, we conducted this work using the traditional method of Likert scaling, resulting in the validation of a 11-item scale. Moving forward, however, we are keen to further investigate, how the items of this trust model relate to each other and how they can be supplemented with other, more innovative methods for quantifying their user response.

The rest of this paper is structured as follows. The second section deals with related works. Section 3 elaborates on the methodology, introducing a new metric scale for measuring trust towards holographic AIs. Section 5 describes discussion and methods to measure values of trust. Finally, conclusions and future work are presented in Section 6.

## **2. RELATED WORK**

Gupta et al., (2019) investigate the sense of trust in virtual reality based on cognitive load levels, physiological sensor data, such as galvanic skin response and heart-rate variability and a system as well as subjective mental effort questionnaire. Although it combines multiple methods and equipment, the processes are complex and difficult to implement.

Interviews, competitive research, and surveys often employ to method human feelings, but it is hard to transform emotional values to numbers or maths. Likert scales are a psychometric instrument used to measure human attitudes and are commonly used in questionnaires. Likert scales apply a trust scale (Borum, 2010) which is used to assign subjective and abstract statements to semantics, translating attitudes, opinions, and feelings to a rated value set in the form of a disagree-agree response scale. The format of five-level items is: 1. Strongly Disagree; 2. Disagree; 3. Undecided; 4. Agree; 5. Strongly agree (Piemetel, 2010).

Likert scales have two parts: statements and response scale. The statement should be a sample, short and clear (Bryman, 2012). Five-point scales, e.g., allow participants to express how much they agree with each statement in order to ensure the explicit differences, interval values and results with no bias. The responses are equated with corresponding integrates, for instance, strongly unfavourable= 1; unfavourable= 2; undecided = 3; favourable = 4; strongly favourable= 5. A high level of agreement with a positive item is endorsed by participants with positive attitudes, and negative perceptions are reverse scored (Kocaballi, Laranjo, and Coiera, 2019). Undecided or neutral decision is able to avoid bias.

For example, in order to analyse the perceptions of interactions with computer assistants, as well as investigate the impressions and experiences of the comprehensibility of the computer agent's communication (including trust), Hanna and Richards (2019) applied Likert scaling to survey degrees of satisfaction with a virtual assistant's verbal conversation. They recruited 73 undergraduate students and tracked the behaviours of all participants, such as inputs and keystrokes. This was done to investigate the relationship between trust, promises, and performance.

Additionally, Kulms and Kopp (2018) also use Likert scales to measure the warmth of computers and how it influences trust. However, they change the strongly agree to disagree format to a trustworthy, good, truthful, well-intentioned, unbiased, and honest format. As a self-reporting measure of trust, Likert scales requires enough data and number of participants to support evidence in order to avoid those extreme cases with biased attitudes. Lack of enough statements with accurate words in a specific context, such as synonyms that have different meanings cannot allow users to make right choices. Observing users' reactions, decisions, and behaviours is able to be an indirect measure of understanding their expectations (Pimentel, 2010).

On the other hand, Thayer proposes a two-dimensional mood model in which valence (happy to unhappy) and degrees of arousal plotted on X and Y axis, describe mood. Multidimensional opinion of emotion is reliable for a single dimensional view (Lang 1980). Self-Assessment Manikin (SAM) is a subjective assessment of emotions based on valence arousal model that applies graphics to judge pleasure, arousal and dominance associated with affective reactions (Bradley and Lang 1994). To indicate arousal states, the images of SAM range from happy (smiling face) to sad (frowning face), excited (wide-eyed image) to calmed (sleepy image). A small picture of a man means minimum control while a large one shows maximum dominance. SAM can be a reference of user experience in order to explore the sense of trust.

## **3. METHODOLOGY: A NEW SCALE FOR 'TRUST'**

There is no validated Likert scale for evaluating trust towards holographic AIs that we consider applicable. We therefore set out to construct a new scale, using Likert scaling (Trochim, 2021) as methodology. We pool items extracted from literature and extended with brainstorming among the researchers first. Subsequently we use a pool of 15 judges to rate the pool of items from 'strongly unfavourable to the concept' to 'strongly favourable to the concept'. Item selection then will be pursued by dropping statements that do not correlate well

with sum scores of all statements. Moreover, via item-item correlations, we can further reduce the number of items by eliminating closely related items. We discard lower values to achieve an administrable scale (e.g.,  $> 0.6$  or  $0.7$ ) that forms a group of optimal statements (Roberts, Laughlin, and Wedell, 1999). Finally, a t-test between top and bottom quarter answers serves to identify those items most suited to polarise answers.

difference in judgement between the judges assigning top scores and the ones assigning lowest scores. This means that items with high t-values are more discriminant, better able to separate. Literature recommends relying on our own judgement in selecting the right trade-off between discriminance and number of items. We chose, based on our analysis, a t-value of over 5.5, resulting in 11 question items (see Figure 6).

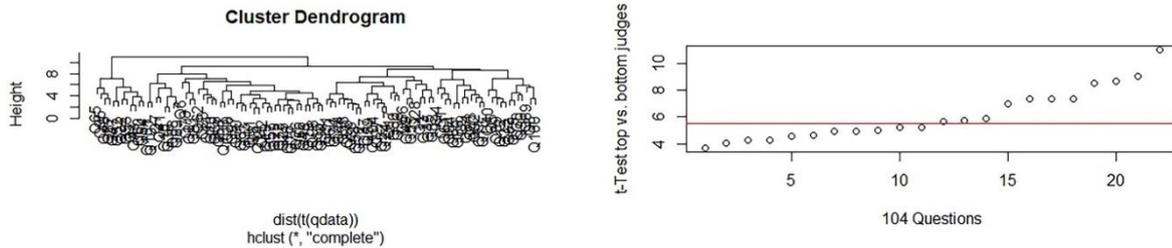


Figure 5. Correlations of the question answer pattern of the judges with the sum scores.

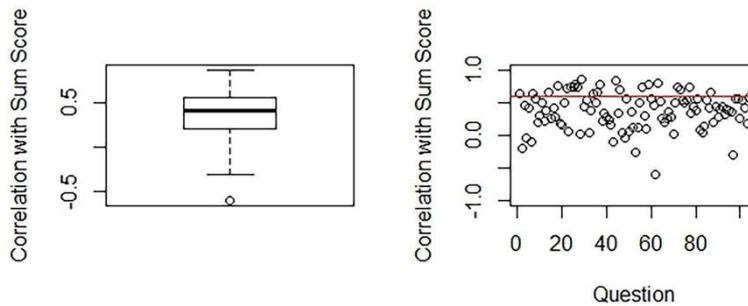


Figure 6. T-values of the top versus bottom quarter judges.

We had identified 104 statements to explore the concept of trust towards holographic AIs, extracted from literature and further extended through brainstorming by the authors. To cut down the selection of items to a reasonable number that can be administered well, we used a pool of judges to explore which of these statements were loading on trust (via the correlation with the sum score) and were polarising.

We eliminate the items where the correlation with the sum scores was less than 0.6, retaining 22 statements (see Figure 5). This removes constructs from the pool, which are not directly loading onto the general direction of all statements – using the direction as a proxy for the core concept of trust.

Next, we calculated the t-test for the mean value of ratings given by the top quarter judges (judges assigning the highest scores) and tested them against the mean values given by the bottom quarter judges (judges assigning the lowest scores), checking whether their t-values indicated strong polarisation. The higher the t-value is, the bigger the

To check, whether we missed any important items, we additionally conducted a cluster analysis over the Euclidean distances of the questions, using hierarchical clustering (*hclust*, see R core team, 2021, package “stats”). We cut the resulting cluster hierarchy in the dendrogram so that questions were clustered into  $k=20$  groups, which from the visual analysis of the cluster dendrogram looked like a good homogeneity level for the resulting clusters.

Subsequently, we manually analysed all resulting groups to systematically double check whether we missed out on particular groups of items. There were several groups that were not retained and looked interesting at first, but when investigating the correlation values of the contained items in the group with the sum scores, it turned out that in all cases, the correlation was too low (below a threshold of 0.6). This clearly indicates that these questions ask about something else, not (or 'not only') about 'trust' (see table 1). The order of items presented in Table 1 has no priority.

We arrived at the conclusion that indeed the statistically motivated selection is the best selection and upon inspection of groupings in the ‘raw’ statements, we can clearly see that other groups measure (also) something else, not just trust.

We pre-tested the questionnaire with the help of a student in a small group (5 participants) (Watts, 2020). All questions are positive statements to ensure a uniform polarity, and this questionnaire with multiple choices, consisting of strongly agree, agree, undecided, strongly disagree and disagree, is able to provide evidence of user experience regarding holographic AIs. Besides, the questionnaire has been verified feasibility. It employed it to investigate the validation of a holographic assistant and an insight of interactions by testing the sense of trust. This questionnaire is able to clearly understand and analyse user intuitionistic feelings and opinions.

## 5. DISCUSSION

**Table 1:** Final metric scale for measuring ‘trust’

#	Item
<i>Competence</i>	
1	The hologram is competent.
2	The hologram is very skilled.
<i>Integrity</i>	
3	I think positively about the hologram
4	The hologram answered my questions truthfully.
<i>Benevolence</i>	
5	I think the hologram wants to do good.
6	The hologram is benevolent.
<i>Compassion</i>	
7	The hologram feels real to me.
8	The hologram looks out for me.
9	The hologram was committed to helping me.
10	The hologram is compassionate.
<i>Relationship</i>	
11	The hologram and I created a good

	relationship.
--	---------------

The final selection retained provides a set of distinct items, all of which contribute to exploring critical elements of trust and explaining whether holographic AIs are trustworthy or not. In line with earlier models published (see introduction), these items can be grouped along constructs as follows.

Statement #3 and #4 (see Table 1) can be interpreted to reflect the aspect of *integrity*, in agreement with Mayer et al.’s proposal (1995).

*Benevolence* can be characterised as the agent’s disposition to do good (Urbano, 2013), grouping #5 and #6 together, following Mayer et al.’s (1995) and Sousa et al.’s (2014) proposals.

*Competence* is instrumental to trust. In the words of McLeod (2020), trust requires that we “rely on others to be competent”. Competence is a potential for human action (Wild, 2016) and when put into practice turns performance. Conversely, the visible demonstration of performance helps establish trust. Competence is typically is defined to subsume skills, knowledge, and abilities (Hager and Gonczi, 2009), and competent agents require a certain degree of domain-specific expertise. Chatbots, for instance, call for communication skills. For intelligent tutors, the pedagogical expertise is key. The agent’s capabilities and their competent application are preconditions to their success. Users consider whether the agents can unsuccessfully achieve their goals and whether this matches with their expectation. The user’s optimism about the agent’s competence (and its subordinate concepts) are assessed through statements #1 and #2.

“The hologram feels real to me” (#7) involves empathy. Empathy is the ability to resonate with other’s emotions. Holographic AIs be emphatic by picking up on affective input of the user and then perform corresponding emotive expressions or other matching behaviour. *Compassion* casts the net a bit wider and has two components, i.e., empathy and the motivation to help (Singer and Klimecki, 2014; Gilbert, 2014). In terms of HCI, compassionate agents not only respond to users’ feelings but are also able to remediate problematic states (Ray, 2018). Compassion refers to both motivation to help and actual helping in response to sensed needs and feelings of others, expressed in statement #8 and #9.

Tseng and Fogg (1999) define trust as “a positive belief about the perceived reliability of, dependability of, and confidence in a person, object, or process”. In our model, trust is an attitude that users have towards holographic AIs based on their belief on whether the agents help achieve goals with good intentions and positive behaviour, ultimately aimed at building a positive relationship with the user.

Mayer et al., (1995) proposed further that interpersonal trust is moderated by a *propensity* to trust, i.e., the willingness to be vulnerable and accept risk based on expectations regarding another person's actions.

Competence could be further assessed using a Turing test (with a Wizard-of-Oz control group agent), providing insights into how the user interface elements influence perception of competence.

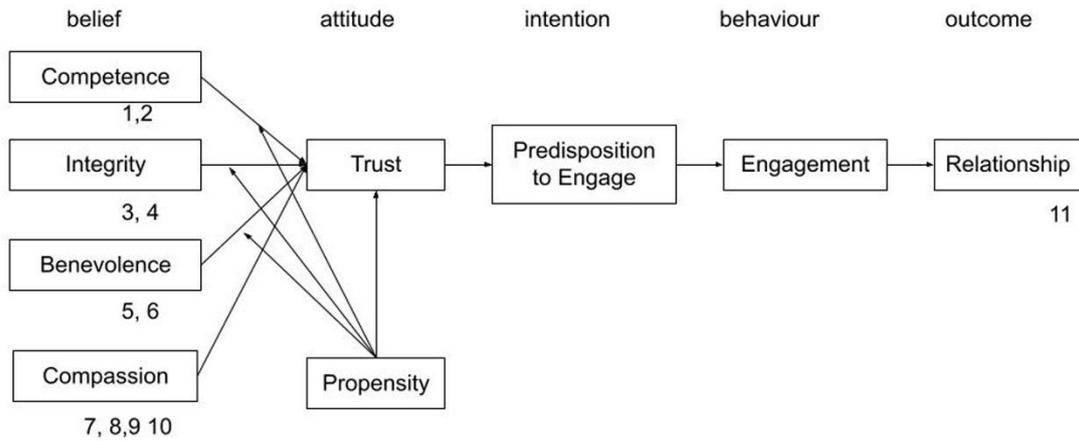


Figure 7. The new model for trust.

We propose that trust is the attitude resulting from the belief assumptions listed above, which then, following Sousa et al. (2014) results on the predisposition to engage, an intension, which ultimately drives engagement behaviour. A measurable result of trust is the establishment of a relationship, as assessed by #11 (see Figure 7).

This could form the basis for a testable structural equation model with trust being a latent variable predicted from the observable ones.

For some of these items, there may also be a more direct approach as alternative to quantify the users' response from Likert scale items. This relates particularly to elements in 'compassion', such as the affective state of the user:

- Sentiment analysis over the dialogue transcripts could provide accurate classifiers detecting affect expressed by the user (and by the AI).
- Facial expressions could provide further insight.
- Prosody of speech can be used to pick up on intonation, as a proxy for affect.
- EEGs are now routinely used to gauge affective state / facial expression of the user in laboratory settings. The imaging technology could offer useful comparison of how the inner, neurological state relates to the more conscious self-assessment expressed in the answers to the questionnaire items.
- Self-assessment manikins (SAMs) can be used as a more graphical replacement, especially useful when dealing with underage participants.

Relationships can be investigated with the help of social network analysis. There are proposals for proxies of engagement in the literature looking at interaction data in the speech dialogue (e.g., interanimation score, see Rebedea et al., 2010).

## 6. CONCLUSION AND FUTURE WORK

In this paper, we reported on the construction of a new metric scale for conceptualising (also quantitatively) an extended model of trust towards holographic AIs, using Likert scaling as development methodology. We identified and refined 104 statements, and then retained eleven of them as a comprehensive model for predicting the level of trust people develop towards holographic AIs.

Trust of holographic AIs is different from interpersonal trust, but also from e-trust towards non-anthropomorphic technology (such as banking apps). Trust is an attitude that such agents help achieve goals, as well as devote themselves to building a positive and interactive relationship with their users. Competence, integrity, benevolence, commitment, and empathy are all key dimensions in this model of trust.

The model provides a questionnaire with Likert scales that we can apply to measure the degree of trust.

In the future, we intend to further investigate whether the questionnaire items of the model can be replaced with other valid measurement methods, like, for example, direct observation. We also would like to investigate, how the questionnaire approach could be emulated with other modalities. For example, we would like to analyse what would change if the questions were asked by the holographic agent directly ("do you find me competent?", "Do I feel real to you?", etc.) and whether that would trigger the same response.

Furthermore, we plan to use self-assessment manikins and a SWOT analysis (focusing on strengths, weakness, opportunities, & threats) to explore user experience and the nature of the relationship between human and agent.

Marinaccio et al., (2015) and Kim and Song (2021) investigated how to recover trust from error. This could be a good idea for our planned pilots, to build in challenging situations in order to derail the relation under stress. Moreover, inappropriate reliance (misuse) and disuse of computer agents may cause distrust (Lee and See, 2004), which we could use to create contrasting behaviour, able to shed more light on the development of trust and how we can influence it.

## ACKNOWLEDGEMENT

We would like to thank 15 anonymous experts to help us fill in the questionnaire for developing this new metric scale. We also would like to thank Rory Watts at Oxford Brookes, who pre-tested the questionnaire for us subsequently. Work on this paper was supported by the European Union's Horizon 2020 research and innovation program through the ARETE project (grant no 856533).

## 7. REFERENCES

- Holz, T. et al. (2011) MiRA—Mixed Reality Agents. In: *International Journal of Human-Computer Studies*, 69(4), 251-268.
- Campbell, A.G., Stafford, J.W., Holz, T., O'Hare, G. (2013) Why, when and how to use augmented reality agents (AuRAs). In: *Virtual Reality* 18(2), 139–159.
- Huang, X., Wild, F., Whitelock, D. (2021) Design Dimensions for Holographic Intelligent Agents: A Comparative Analysis. AIED'2021, Netherlands.
- Pinxteren, M.M., Pluymaekers, M., & Lemmink, J. (2020) Human-like communication in conversational agents: a literature review and research agenda. *Journal of Service Management*, 31, 203-225.
- Mayer, R. C., Davis, J. H., & Schoorman, F. D. (1995) An integrative model of organizational trust. *Academy of Management Review*. 20, 709-734.
- Nass, C., Steuer, J., and Tauber, E. R. (1994) Computers are social actors. CHI '94 Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, New York, NY: ACM.
- Dobel, J.P. (1999) *Public Integrity*. John Hopkins University Press, Baltimore.
- Philip, P., Dupuy, L., Auriacombe, M., Serre, F., Sevin, É.D., Sauteraud, A., & Micoulaud-Franchi, J. (2020) Trust and acceptance of a virtual psychiatric interview between embodied conversational agents and outpatients. *NPJ Digital Medicine*, 3.
- Borum, R. (2010) *The Science of Interpersonal Trust*. Mental Health Law & Policy Faculty Publications. 574.
- Sousa, S., Lamas, D., & Dias, P. (2014) A Model for Human-Computer Trust: Contributions Towards Leveraging User Engagement. *Learning and Collaboration Technologies*, Greece, June 22, 128-137. Springer, Cham.
- Jones, G., & George, J. (1998). The experience and evolution of trust: Implications for cooperation and teamwork. *Academy of Management Review*, 23, 531-546.
- Kulms, P., & Kopp, S. (2018) A Social Cognition Perspective on Human-Computer Trust: The Effect of Perceived Warmth and Competence on Trust in Decision-Making With Computers. *Frontiers Digit. Humanit.*, 5, 14.
- Lee, J. J., Knox, W. B., Wormwood, J. B., Breazeal, C., and DeSteno, D. (2013) Computationally modeling interpersonal trust. *Front. Psychol.* 4, 893.
- Lee, J., & See, K.A. (2004) Trust in Automation: Designing for Appropriate Reliance. *Human factors and Ergonomics society*, 46 1, 50-80.
- Gupta, K., Hajjika, R., Pai, Y.S., Duenser, A., Lochner, M., & Billinghamurst, M. (2019) In AI We Trust: Investigating the Relationship between Biosignals, Trust and Cognitive Load in VR. 25th ACM Symposium on Virtual Reality Software and Technology, New York, November 12, 1-10.
- Pimentel, J. (2010) A note on the usage of Likert Scaling for research data analysis. *USM R&D Journal*, 18, 109-112.
- Hanna, N., Richards, D. (2019) Speech Act Theory as an Evaluation Tool for Human-Agent Communication. *Algorithms*. 12. 79.
- Lang, P.J. (1980) Behavioural treatment and bio-behavioural assessment: computer applications, *Technology in mental health care delivery systems*, 119-137.
- Bradley, M.M. and Lang, P.J. (1994) Measuring emotion: The self-assessment Manikin and the semantic differential. *Journal of Behaviour Therapy and Experimental Psychiatry*, 25, 49-59.
- Trochim, W.M.K. (2021): Likert Scaling, Conjoint.ly, online at <https://conjointly.com/kb/likert-scaling/>, last access 06/2021.

- Bryman, A. (2012) *Social Research Methods*. OUP Oxford.
- Kocaballi, A. B., Laranjo, L., and Coiera, E. (2019) Understanding and measuring user experience in conversational interfaces. *Interacting with Computers*, 31, 192–207.
- Roberts, J. S., Laughlin, J. E., & Wedell, D. H. (1999) Validity issues in the Likert and Thurstone approaches to attitude measurement. *Educational and Psychological Measurement*, 59, 211–233.
- R Core Team (2021) *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>.
- Watts, R. (2020) An Interactive, Holographic 3D Model using IBM Watson.
- Urbano, J., Rocha, A.P., Oliveira, E. (2013) The impact of benevolence in computational trust. *Agreement Technologies*, Beijing, August 1, 210–224. Springer, Heidelberg.
- Ray, A. (2018) *Compassionate Artificial Intelligence: Frameworks and Algorithms*. Compassionate AI Lab.
- Colquitt, J. A., Scott, B. A., and LePine, J. A. (2007) Trust, trustworthiness, and trust propensity: a meta-analytic test of their unique relationships with risk taking and job performance. *J. Appl. Psychol.* 92, 909–927.
- Marinaccio, K., Kohn, S., Parasuraman, R., & De Visser, E. J. (2015) A Framework for Rebuilding Trust in Social Automation Across Health-Care Domains. *Proceedings of the International Symposium on Human Factors and Ergonomics in Health Care*, 4, 201–205.
- Kim, T., & Song, H. (2021) How should intelligent agents apologize to restore trust?: The interaction effect between anthropomorphism and apology attribution on trust repair. *Telematics and Informatics*, 61, 101595.
- McLeod, C. (2020) Trust. In: Zalta (Ed.): *The Stanford Encyclopedia of Philosophy* (substantive revision Mon Aug 10, 2020), Stanford University, ISSN 1095-5054.
- Wild, F. (2016) Learning Theory and Algorithmic Quality Characteristics. In: *Learning Analytics in R with SNA, LSA, and MPIA*. Springer, 23-44.
- Hager, P., Gonczi, A. (2009) What is competence? *Medical teacher* 18, 15-18.
- Legg, S., Hutter, M. (2006) A formal measure of machine intelligence. In *Proc. 15th Annual Machine Learning Conference of Belgium and The Netherlands (Benelearn'06)*, Ghent, 73–80.
- Fiske, S.T., Cuddy, A.J., Glick, P. (2007) Universal dimensions of social cognition: warmth and competence. *Trends in Cognitive Sciences* 11(2), 77–83.
- Singer, T., Klimecki, O. (2014) Empathy and compassion, *Current Biology*, 24, 875-878.
- Gilbert, P. (2014) The origins and nature of compassion focused therapy. *Br. J. Clin. Psychol.* 53, 6–41.
- Rebedea, T., Dascalu, M., Trausan-Matu, S., Banica, D., Gartner, A., Chiru, C., Mihaila, D. (2010) Overview and Preliminary Results of Using PolyCAFe for Collaboration Analysis and Feedback Generation. In: Wolpers, Kirschner, Scheffel, Lindstaedt, Dimitrova (Eds.): *Sustaining TEL: From Innovation to Learning and Practice. EC-TEL 2010. Lecture Notes in Computer Science*, vol 6383. Springer, Berlin, Heidelberg. [https://doi.org/10.1007/978-3-642-16020-2\\_31](https://doi.org/10.1007/978-3-642-16020-2_31)