World Academy of Science, Engineering and Technology
International Journal of Computer and Information Engineering
Vol:15, No:2, 2021

Trusting Smart Speakers: Analysing the Different Levels of Trust between Technologies

Alec Wells, Aminu Bello Usman, Justin McKeown

Abstract**—**The growing usage of smart speakers raises many privacy and trust concerns compared to other technologies such as smart phones and computers. In this study, a proxy measure of trust is used to gauge users' opinions on three different technologies based on an empirical study, and to understand which technology most people are most likely to trust. The collected data were analysed using the Kruskal-Wallis H test to determine the statistical differences between the users' trust level of the three technologies: smart speaker, computer and smart phone. The findings of the study revealed that despite the wide acceptance, ease of use and reputation of smart speakers, people find it difficult to trust smart speakers with their sensitive information via the Direct Voice Input (DVI) and would prefer to use a keyboard or touchscreen offered by computers and smart phones. Findings from this study can inform future work on users' trust in technology based on perceived ease of use, reputation, perceived credibility and risk of using technologies via DVI.

Keywords**—**Direct voice input, risk, security, technology and trust.

I. Introduction

AS new technologies are constantly emerging, it is important to understand how the user perceives such technologies, especially when considering possible security concerns. For several years, studies in information security and practice have posed, that many security incidents and breaches are caused by human factors, rather than technical failures [1]. Thus, human factors on the part of individual or an employee in an organisation can be an attack vector that can be exploited. The assessment of human factors as an attack vector is a complex multi-component and multi-level problem involving characteristics of hardware, software, user interface, and how humans' issue instructions to technologies [2]. The use of DVI to issue instructions to technologies is growing rapidly and it appeared to be part of the future of human-computer interaction [3]. DVI has provided what may be a more 'natural' mode of control for communication between human and technology that is more convenient and faster compared with the conventional input means; such as the keyboards and touch screens [4]. Although, DVI enables novel and convenient forms of interaction with technologies, there are concerns about the issues of trust, security and privacy on using DVI [3].

At the heart of human-centred security assessments, trust (soft security property) [5] is considered an important social control security mechanism which can be used to evaluate the security aptitude of human agents. The trust mechanism supplements traditional technical hard security mechanisms (e.g., authentication and access control), thus enables a wider view of social control mechanisms in addition to existing technical-based security approaches towards an overall cyber security system. Trust is seen as a dynamic concept that can constantly change in humans due to its many different facets and dimensions. Studies on trust in e-commerce have focused on users' responses to interface design and the complexity rather than the deeper, emotion-charged dynamic of trust [6] whereas other studies on trust and digital system focused on evaluating trust based on usability, perceived privacy, and content requirements [7]. In relation to technologies, trust can be considered as the degree to which a user believes in the veracity or effectiveness of a technology to function expectedly either based on the credibility, reputability of the technology, or simply based on the users' experience and perceptions [8]. Thus, in this context, trust can be viewed as the confidence a user can have about the device to behave in an expected manner [9].

Technology is increasingly becoming a part of human lives and is considered an extension of human functions [8]. As a society today, we treat digital technology tools and algorithms with extensive trust. The extent and the degree of how we trust technologies with sensitive data and our lives will soon be a security concern, if it is not already a concern. In other words, we constantly share our sensitive data, delegate responsibilities to technologies and apparently, we trust them. This seems the case particularly in the modern world where new technologies are often being developed and quickly adopted by homes and workplaces. For example, the adoption of smart speakers is already in 13% of United States households and in the United Kingdom, about 10% of households have already adopted Smart Speakers. This was also projected to grow to 55% in the United States and 48% in the United Kingdom by 2022 [10]. Subsequently, based on the increase of people adoption and reliance on these technologies, it can be understood that the more we adopt these technologies, the more urgent becomes the issue of security, trust and associated risk [11]. However, the details of security models and algorithms (e.g. the encryption or the security cloud architecture) used in these technologies may not always be black and white, or a concept that can easily be understood, especially for a non-technical user. Thus, the usable security measure of those devices in the eyes of a non-technical user can come down to – a very simple binary condition; trust or no trust.

Alec Wells, Aminu Bello Usman, and Justin McKeown are with the Department of Computer Science, York St John University, York, UK (e-mail: alec.wells@hotmail.co.uk, a.usman@yorksj.ac.uk, J.McKeown@yorksj.ac.uk).

The study is interested in understanding whether users have varying degrees of trust depending on the technology's input method, DVI, keyboard or touch screen, with the hope to improve understanding of security concerns when adopting DVI in technologies. It is also important to understand whether the users are more likely to give away sensitive information because of the trust they have with a technology and the input method used to access the technology. The study considers three different technologies; *smart speakers*, *smart phones* and *computers* and compares the participants' opinions about trust between each of the technologies. This will hopefully identify if there are any security concerns with adopting voice-based technologies such as smart speakers.

The contribution of the paper can be summarised as follows.

- Discuss the current state of trust and technology in relation to smart speakers, smart phones and computers.
- Investigate if there are any differences in how people perceive the technologies and if they would be more likely to trust them with sensitive information.
- Evaluate if people have more or less trust with certain technologies depending on the input method.

The paper is structured as follows: Section II looks at the related work, with the concept of trust in social context as well as trust and technology with emphasis on the three technologies (smart speakers, smart phones and computers). Section III presents the trust model. Section IV discusses the research methodology and the description of the precipitants involved in the study. Section V presents the results of the study and discussion. In Section VI, a conclusion is presented along with potential future work.

II. Related Work

Within the modern world, as more of our lives become dependent on digital technologies, it is an ever-growing concern for people to keep that part of our lives secure and private. This can be true from the perspective of average users who want to keep private or sensitive information away from prying eyes, or from those with malicious intent. This growing awareness for keeping data secure has only become more apparent as new information and scandals come to light. One of the recent examples of this might come from the Facebook-Cambridge Analytica data scandal. This was a huge breach in security for many users, in which their data were collected and used to influence political adverts to have the most impact on those specific users [12]. One possible explanation why user data and privacy can easily be breached; leak or be compromised is due to how we use and trust those technologies. As new technologies are invented, so are new ways of interacting with technologies, like DVI, and this subsequently results in rendering new challenges about how we can trust the technologies.

A. Trust in Social Context

The concept of trust is usually applied in the context of social relationships between social agents, which can be defined as a social construct with natural attributes to relationship between social actors (a group or individual). Trust can also be viewed as being subjective and a unidirectional relation between social agents and how social agent assess another agent or groups to perform a particular action with a certain level of probability [13]. Simply, trust can be attributed to relationships between people and attributes, trust is subjective, dynamic, and it can evolve with time, experience and the environment. Uslaner et al. describes trust as "*the chicken soup of social life*" [14] – it works mysteriously, often, and we develop trust with only people we know, yet the benefits of trust mostly come from when we trust strangers. For example, a service provider and customers, where a customer does not know the service provider, yet the customers can trust the service provider with their personal and sometimes banking details.

B. Trust and Technology

Arguably, one similarity between trust in the context of social relations and trust between humans and technology, is that humans can develop trust with technology that they have found to be comfortable using, a technology can make them feel safe either due to its functionalities, reputation or its perceived credibility. However, trust with technologies takes time to establish and mostly, users trust technology that they have found to be more reliable over time and only the individual has actual perceptions about how much they trust the technology.

Trust is similarly described in the context of human-robot interaction as humans must trust that a robotic teammate will protect the interests and welfare of every other individual on the team. For users to gain the advantages and benefits of robotic teammates, they must be willing to trust and accept robot-produced information and follow their suggestions [15]. Much like in trust with social agents and companies, for oneself to benefit from robots, trust must first be established. In addition, throughout trials in an experiment it is interesting to note that levels of trust change over time, based upon the reliability of the automation. This was observed in pilots that constantly used automation, who were found to trust automation more often than students [16]. Hence, initial views can be different to those observed later. When measuring trust in human-robot collaboration, studies utilise a performance model and observe different research areas such as psychology behind team performance, unmanned systems, mixed initiative systems and war fighting behaviour which is adapted when identifying how much humans trust robots' decisions [17]. However, it is important to consider that that trust between humans and autonomous agents can be different between trust with humans and other such things due to the fact autonomous agents are not human and trust with autonomous machines largely is about trusting that the machine will perform as intended [18].

III. Trust Model

In this study, Corritore et al. model is used to develop the proposed questions and the basis of the approach, due to the models continued relevance in representing how trust is

formed [19]. The model also supports the literature above, about how trust can be formed with technologies from a user's perception of factors. The model can be seen in Fig. 1 with three perceptual factors that impact on trust, namely: perception of credibility, ease of use and risk. As shown in Fig. 1, the model identifies two categories that can contribute to building an individual trust about a particular technology; perceived factors and external factors. The external factors of a technology affect the perceived factors a user has about a technology. Some examples of external factors include the experience a user has with the technology, the devices portability and the control the user has in interacting with said technology. The external factors comprise the physical and psychological factors surrounding a specific technology. Perceived factors fall into the following three categories: easiness of use, credibility, and potential risk, described as follows.

- Perception of ease of use reflects the degree to which a person believes that using a technology would be free of effort [20]. Ease of use can be separated into two categories; how easy it is to learn and how easy it is to use.
- Perception of credibility comprises the following dimensions: believability, integrity, reputability, vulnerability, advantage and hostility [21].
- Perception of risk can be viewed as how the users perceive risk when the security of their devices for securing their personal information is not verified [22].

As illustrated in the presented model of Fig. 1, the relationships between the model's elements are external factors to perceived factors, perceived factors to other perceived factors and perceived factors to trust. To this point, it can be inferred that the perceived ease of use of technology, the users' perception about the credibility of the technology and associated risk can contribute to determining how users can trust a device with sensitive data. These features can be defined as external factors. Based on the above about Fig. 1, it can be deduced that the level of trust a user can have regarding a technology can directly be linked to how credible a device is and the potential risk of using it in the eyes of the user. Based on these factors a proxy measure of trust can be utilised to evaluate the trust via a question based on the presented model in Fig. 1.

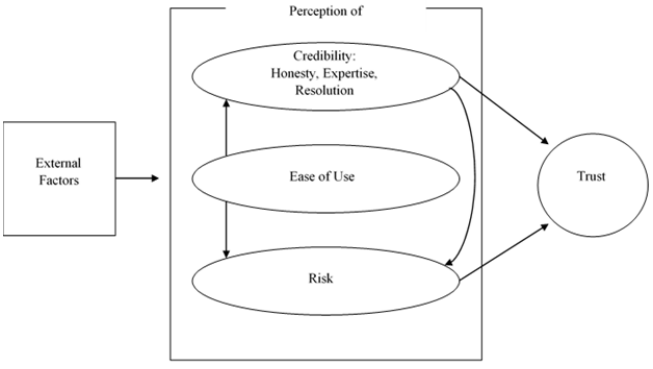

Fig. 1 Trust Model

IV. Methodology

For the experiment, a between-groups testing method was used rather than a repeated measures method. This means that participants were separated into 3 different groups, with each group only testing 1 technology. In comparison, a repeated measure would involve participants utilising all 3 technologies. Primarily, this was done to improve the accuracy of the results as participants will not be influenced by their answers given when utilising the other technologies. The dependent variable of this study is the measurement of trust, which is measured via the questionnaire participants filled after using the technology; in contrast, the independent variable is the 3 technologies that participants will interact with. Three different technologies were used: Amazon Echo Dot Smart Speaker, a Sony Xperia smart phone running android and a standard university computer running Windows 10 operating system. Each technology was used by separate control groups. Participants were first briefed that they would be asked to individually fill out a sign-up sheet that would emulate a sign-up process to a website, using the technology they were assigned and would have to give personal information such as their names, emails and setting a passwords. They were then given a consent form and took part in the study. Both the computer and smart phone sign-up sheets were designed to be as plain as possible, on a white background with only answer boxes and labels, as to not influence participants perceptions by including design features like logos as that could affect how credible some users deem the technology to be. For the smart speaker, a chatbot from Bot Libre was utilised to ask the same fields asked in the computer/smart phone version, meaning the smart speaker itself was not actually asking question, though it appeared to be. This is known as the 'wizard of oz' technique, which was done to provide the exact same questions across all 3 groups.

18 participants took part in this study, out of which 11 participants were male and 7 were female. The ages of participants ranged between 20-60, and all participants were randomly assigned one of the 3 technologies (smart speaker, smart phone and computer). After using the technologies to give sensitive information via a sign-up sheet, participants were then asked to fill out a questionnaire, whose questions are based on the trust model of Fig. 1 to obtain a proxy measure of trust as described in the following section.

V. Results Analysis and Discussion

A. Result Analysis

To analyse the results, the data were examined (the users' responses to the survey) using a rank-based nonparametric test known as the 'Kruskal-Wallis H Test' to determine if there was a statistically significant differences between the users' trust level and the independent variables (smart speaker, computer and phone) between each of the groups.

The results of the analysis can be seen in the appended Table I. The table shows the mean rank, standard deviation, Kruskal-Wallis score and the assumed significant figure for each question that participants were asked. Each question was

ranked on a scale of 1-5 with 1 being strongly disagree and 5 being strongly agree. As mentioned earlier, 18 participants in total took part in the experiment and were divided by technology; hence each technology had responses from 6 different participants. The Kruskal-Wallis value is found by distributing chi-squared, which is then used to determine the p value and verify if the data is statistically significant. The mean value indicates which group had higher scores. A higher score would be better for all questions except when asked if participants felt vulnerable, cautious or that it was Risky to use the technology in which case a lower score is better.

The results show that when asked if "The technologies information required is **believable**?", there was no statistical significance between the technologies, X^2 (2) = 2.807, p = 0.246. Though in terms of mean ranking score, computer was the highest with 12.08, speaker was the second with 8.58 and smart phone with the least at 7.83.

The results show that when asked if "The technology has **integrity**?", there was no statistical significance between the technologies, X^2 (2) = 0.506, p = 0.776. Though in terms of mean ranking score, computer was the highest with 10.17, speaker was the second with 10.00 and smart phone with the least at 9.33.

The results show that that when asked if "The technology is **reputable**?", there was no statistical significance between the technologies, X^2 (2) = 0.273, p = 0.873. Though in terms of mean ranking score, speaker was the highest with 10.33, smart phone was the second with 9.33 and computer with the least at 8.83.

The results show that when asked if "The technology is **respected?**", there was no statistical significance between the technologies, X^2 (2) = 0.505, p = 0.777. Though in terms of mean ranking score, computer and smart phone were tied the highest at 10.08 and speaker the lowest at 8.33.

The results show that when asked if "The technology was what I **expected**?", there was no statistical significance between the technologies, X^2 (2) = 0.711, p = 0.701. Though in terms of mean ranking score, computer was the highest with 10.83 with smart phone and speaker tied at the lowest with 8.83.

The results show that when asked if "The technology was **predictable**?", there was no statistical significance between the technologies, X^2 (2) = 4.229, p = 0.121. Though in terms of mean ranking score, computer was the highest with 12.00, smart phone was the second with 9.50 and speaker with the least at 7.00.

The results show that when asked if "**Learning to use** the technology was easy?", there was no statistical significance between the technologies, X^2 (2) = 4.192, p = 0.123. Though in terms of mean ranking score, computer was the highest with 12.50, smart phone was the second with 8.25 and speaker with the least at 7.75.

The results show that when asked if "I found the technology **easy to use**?", there was no statistical significance between the technologies, X^2 (2) = 3.490, p = 0.175. Though in terms of mean ranking score, computer was the highest with 12.50, smart phone was the second with 8.92 and speaker with the least at 7.58.

The results show that when asked if "I felt **vulnerable** using the technology?", there was no statistical significance between the technologies, X^2 (2) = 5.395, p = 0.067. Though in terms of mean ranking score, speaker was the highest with 13.33, computer was the second with 8.50 and smart phone with the least at 6.67.

The results show that when asked if "I feel like I must be **cautious** using the technology?", there was no statistical significance between the technologies, X^2 (2) = 0.022, p = 0.989. Though in terms of mean ranking score, smart phone was the highest with 9.75, computer was the second with 9.42 and speaker with the least at 9.33.

The results show that when asked if "It is **risky** to use the technology?", there was no statistical significance between the technologies, X^2 (2) = 1.236, p = 0.539. Though in terms of mean ranking score, speaker was the highest with 10.58, smart phone was the second with 10.33 and computer with the least at 7.58.

The results show that when asked if "I believe the technology won't take **advantage of me**?", there was no statistical significance between the technologies, X^2 (2) = 3.983, p = 0.137. Though in terms of mean ranking score, smart phone was the highest with 11.75, computer was the second with 10.67 and speaker with the least at 6.08.

The results show that when asked if "I believe the technology will not use my data **maliciously**?", there was no statistical significance between the technologies, X2 (2) = 4.684, p = 0.096. Though in terms of mean ranking score, mobile was the highest with 12.00, computer was the second with 10.67 and speaker with the least at 5.83.

IV. Conclusion

In this study, Corritore's et al. model [19] is is used to formulate a proxy measure of trust with three different technologies: smart speakers, smartphones, and computers. The purpose was to understand which technology most people were likely to trust. Kruskal-Wallis H test was used to examine the collected data and determine if there was a statistical significance on the level of trust the users have about the three technologies. The study found that participants preferred to use a keyboard for a computer or touch screen for a smart phone over using DVI when determining which technology to trust when handling their sensitive data. Despite this however, many participants considered that DVI was extremely easy to use, but they also felt it was one of the most reputable and widely accepted of the 3 technologies. In terms of practical implications of the study, it may be a consideration to those developing voice-based technologies, that people may have security concerns with the devices. On the other hand, it is not yet known about the long-term effects, since the smart speakers and other voice-based technologies (such as voice-based assistants in smart phones) are still relatively new, hence in a few years, perceptions of voice technologies may have changed. In the next phase of this study, a possible consideration could be looking at technologies being used where levels of trust were obtained

over a long period of time. Likewise, as new technologies are developed, people could perceive voice-based technology in a different light and may produce wildly different results.

APPENDIX

TABLE I
RESULT SUMMARY

Question	Technology	Mean Rank	Standard Deviation	Kruskal-Wallis	Asymp. Sig.	Number of Participants
The technologies information required is **believable**?	Speaker	8.58	0.6183	2.807	0.246	6
	Computer	12.08				6
	Smart phone	7.83				6
The technology has **integrity**?	Speaker	10.00	1.0226	0.506	0.776	6
	Computer	10.17				6
	Smart phone	8.33				6
The technology is **reputable**?	Speaker	10.33	1.1100	0.273	0.873	6
	Computer	8.83				6
	Smart phone	9.33				6
The technology is **respected**?	Speaker	8.33	0.9376	0.505	0.777	6
	Computer	10.08				6
	Smart phone	10.08				6
The technology was what I **expected**?	Speaker	8.83	0.6077	0.711	0.701	6
	Computer	10.83				6
	Smart phone	8.83				6
The technology was **predictable**?	Speaker	7.00	1.0416	4.229	0.121	6
	Computer	12.00				6
	Smart phone	9.50				6
Learning to use the technology was easy?	Speaker	7.75	0.6077	4.192	0.123	6
	Computer	12.50				6
	Smart phone	8.25				6
I found the technology **easy to use**?	Speaker	7.58	0.6978	3.490	0.175	6
	Computer	12.50				6
	Smart phone	8.92				6
I felt **vulnerable** using the technology?	Speaker	13.33	1.5424	5.395	0.067	6
	Computer	8.50				6
	Smart phone	6.67				6
I feel like I must be **cautious** using the technology?	Speaker	9.33	1.4642	0.022	0.989	6
	Computer	9.42				6
	Smart phone	9.75				6
It is **risky** to use the technology?	Speaker	10.58	1.3921	1.236	0.539	6
	Computer	7.58				6
	Smart phone	10.33				6
I believe the technology won't take **advantage** of me?	Speaker	6.08	1.3492	3.983	0.137	6
	Computer	10.67				6
	Smart phone	11.75				6
I believe the technology will not use my data **maliciously**?	Speaker	5.83	1.3394	4.684	0.096	6
	Computer	10.67				6
	Smart phone	12.00				6

REFERENCES

[1] S. Bruce, "Secrets and Lies–Digital Security in a Networked World," 2000.
[2] D. Henshel et al, "Trust as a human factor in holistic cyber security risk assessment," Procedia Manufacturing, vol. 3, pp. 1117-1124, 2015.
[3] Nic Newman, "The future of voice and the implications for news " Reuters Institute for the Study of Journalism, United Kingdom, November. 2018.
[4] P. R. Cohen and S. L. Oviatt, "The role of voice input for human-machine communication," Proceedings of the National Academy of Sciences, vol. 92, (22), pp. 9921-9927, 1995.
[5] P. Ratnasingham, "The importance of trust in electronic commerce," Internet Research, vol. 8, (4), pp. 313-321, 1998.
[6] P. L. Hardré, "When, how, and why do we trust technology too much?" in Emotions, Technology, and Behaviors Anonymous 2016.
[7] M. A. Sasse, "Usability and trust in information systems," in Anonymous 2005.
[8] A. H. Kiran and P. Verbeek, "Trusting ourselves to technology," Knowledge, Technology & Policy, vol. 23, (3-4), pp. 409-427, 2010.
[9] W. Sherchan, S. Nepal and C. Paris, "A survey of trust in social networks," ACM Computing Surveys (CSUR), vol. 45, (4), pp. 47, 2013.
[10] (September, 5). Voice-enabled smart speakers to reach 55% of U.S. households by 2022, says report. Available: https://techcrunch.com/2017/11/08/voice-enabled-smart-speakers-to-reach-55-of-u-s-households-by-2022-says-report/.
[11] M. Coeckelbergh, "Can we trust robots?" Ethics and Information Technology, vol. 14, (1), pp. 53-60, 2012.
[12] C. Cadwalladr and E. Graham-Harrison, "Revealed: 50 million Facebook profiles harvested for Cambridge Analytica in major data breach," The Guardian, vol. 17, pp. 2018, 2018.
[13] A. B. Usman and J. Gutierrez, "DATM: a dynamic attribute trust

model for efficient collaborative routing," Annals of Operations Research, vol. 277, (2), pp. 293-310, 2019.
[14] E. Uslaner, "The Moral Foundations of Trust University of Maryland," College Park, MD, 1999.
[15] P. A. Hancock et al, "A meta-analysis of factors affecting trust in human-robot interaction," Hum. Factors, vol. 53, (5), pp. 517-527, 2011.
[16] C. Gold et al, "Trust in automation–Before and after the experience of take-over scenarios in a highly automated vehicle," Procedia Manufacturing, vol. 3, pp. 3025-3032, 2015.
[17] A. Freedy et al, "Measurement of trust in human-robot collaboration," in 2007 International Symposium on Collaborative Technologies and Systems, 2007, .
[18] D. J. Atkinson and M. H. Clark, "Autonomous agents and human interpersonal trust: Can we engineer a human-machine social interface for trust?" in 2013 AAAI Spring Symposium Series, 2013.
[19] C. L. Corritore, B. Kracher and S. Wiedenbeck, "On-line trust: concepts, evolving themes, a model," International Journal of Human-Computer Studies, vol. 58, (6), pp. 737-758, 2003.
[20] Y. He, Q. Chen and S. Kitkuakul, "Regulatory focus and technology acceptance: Perceived ease of use and usefulness as efficacy," Cogent Business & Management, vol. 5, (1), pp. 1459006, 2018.
[21] S. Tseng and B. J. Fogg, "Credibility and computing technology," Commun ACM, vol. 42, (5), pp. 39-44, 1999.
[22] R. Schnall et al, "Trust, perceived risk, perceived ease of use and perceived usefulness as factors related to mHealth technology use," Stud. Health Technol. Inform., vol. 216, pp. 467, 2015.