

How Trust is Defined and its use in Human-Human and Human-Machine Interaction

Kaplan, A.D., Kessler, T.T., and Hancock, P.A.

Trust is a critical element in virtually all forms of interactions, including those between humans and machines. Yet aspects of trust do vary somewhat between definitions. The present work seeks to unify these disparate definitions of trust, comparing and contrasting between the major works. Overall, every definition of trust involves an individual in a position of vulnerability (the trustor) and a person on whom they must rely (the trustee) despite circumstances which may place the trustor in some kind of potential for harm. Such engagements are enacted in order to secure some form of gain from the trusting relationship. The ways in which these definitions influence empirical measurement (both qualitative and quantitative) are identified and elaborated on.

INTRODUCTION

Trust is imperative to properly functioning relationships between a human and another entity. Within human factors and ergonomics, interaction partners can be human, automation, or robot (Hancock et al., 2011; Schaefer et al., 2016). However, at present there is no one single agreed-upon definition of trust for any single of these entities, let alone for a definition that covers them as a whole. As a result of such extant disparities, the scales, surveys, and qualitative measures used to assess trust necessarily differ both between disciplines and within fields of study. This makes efforts to compare findings across disciplines difficult, if not invalid. To answer the question, ‘What is Trust?’ we present a synthesized summary of definitions in order to distill their commonalities and describe their varying differences. In answering that question, one discovers not simply how trust is described, but how it is viewed across fields, how it is measured, and how it is employed in real world situations.

DEFINITIONS OF TRUST

Though trust is defined differently in multiple cases, in all instances of trust there are three fundamental elements: 1) a trustor, who is doing the trusting and who is vulnerable to harm from another individual. 2) the trustee, who is the one being trusted, and who is the individual capable of harming the trustor; and, 3) a context within which the trustee's actions are capable of causing harm or benefit to the trustor. Commonly-cited definitions of trust are listed in Table 1.

SIMILARITIES

A word-cloud was generated from the definitions of trust listed in Table 1(see Figure 1). This cloud gives the highest weight, and thus largest font size,

to words which appear more frequently in the definitions of trust. The lighter, smaller words are those which are used less extensively in multiple definitions. Words that were essentially the same, such as “vulnerable” and “vulnerability” were combined, as were different tenses or spellings of the same word.



Figure 1: A generated word cloud concerning the current definitions of trust

It is clear that, despite “trust” being the product of an interaction, the word “individual” is among the most frequently used. Also highly represented are the words “vulnerable,” “expectation,” “party,” and “action.” Even though there is a fair amount of variation in the definitions, the word cloud shows that there are some rather clear commonalities.

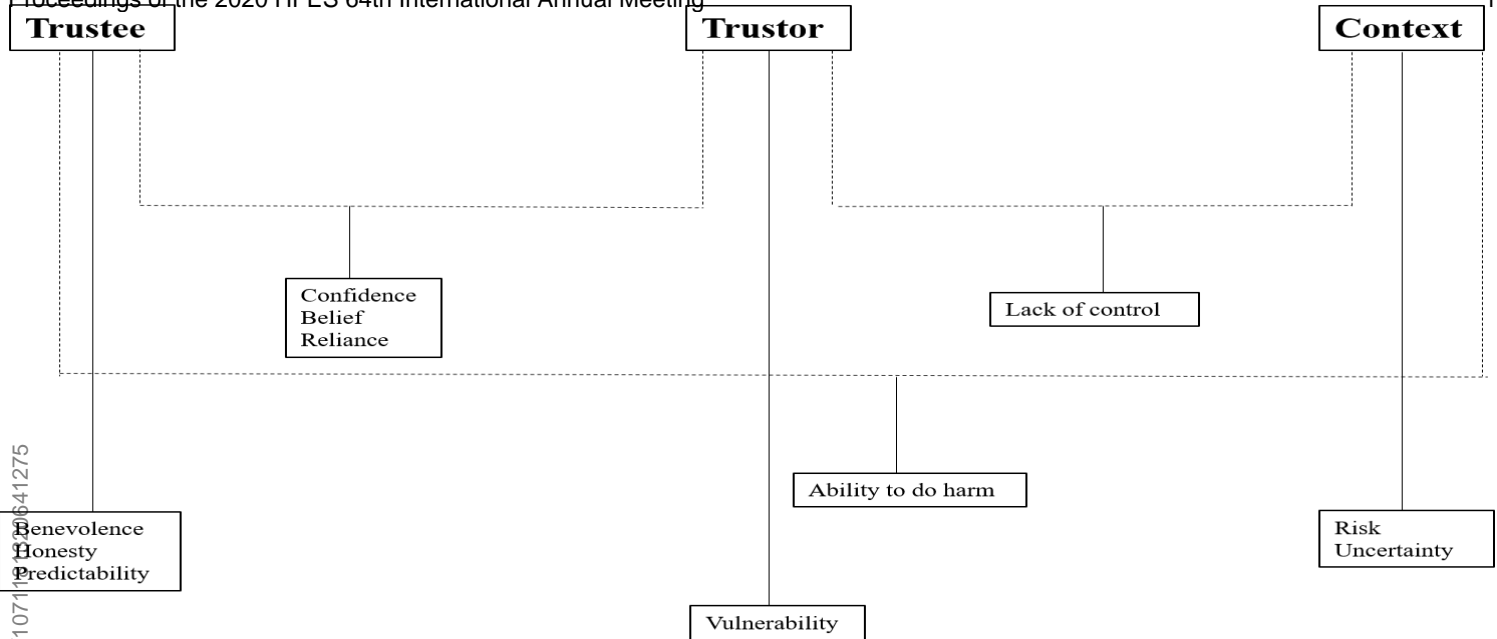


Figure 2: A model of the definitions of trust, separated by category.

DIFFERENCES

Some definitions of trust center squarely on the trustee. Larzelere and Huston (1980) define trust entirely based on perceptions of the trustee- that he is benevolent and honest, similar to Rotter's assertion (1967) that trust exists when the trustee is reliable. McAllister (1995) considers trust to be a trustor's confidence that the trustee is competent and considerate. Gambetta (1988) and Cummings and Bromiley (1996) also center trust around the trustee's actions, with Gambetta stating that trust is the probability of the trustee behaving in a specified manner, with Cummings and Bromily defining a trustworthy individual as one who is honest, makes good faith efforts, and does not take advantage. Zaheer, McEvily, and Perrone (1998) note that one is trustworthy when they are reliable and predictable, even when being dishonest could benefit them. Mishra (1996) claims that trust is the belief that the trustee will be competent, open, concerned, and reliable. Hoy and Tschannen-Moran (1999) also mention the trustor's benevolence, reliability, competence, honesty, and openness. While these definitions vary somewhat, they all focus primarily on the traits of the trustee, namely reliability and competence. ISA/IEC (2011) defines trust as the belief that the trustee will act as intended. Hancock et al. (2011) describe trust not as an expectation of how a trustee will behave, but what actions the trustee will *not* take, and call trust a belief by the trustor that another will not actively seek to harm them.

Others have defined trust primarily in terms of the trustor's vulnerability. Mayer, Davis, and Schoorman (1995), for example, noted that trust was a situation

when the trustor is willing to be vulnerable to another, even when they have no control over the outcome of a situation. Zand (1997) called trust the increase of vulnerability in instances where one can be abused. Rousseau and colleagues (1998) considered trust to be a psychological state on the part of the trustor, who is willing to be vulnerable to the trustee. Several of the previously mentioned authors also took note of the trustor's vulnerability. Mayer, Davis, and Schoorman (1995), Mishra (1996), and Hoy and Tschannen-Moran (1999) all referenced the trustor's susceptibility to harm at the hands of the trustee.

Some definitions focus on the shared contextual environment, namely the risky nature of many situations where trust exists. Lee and See (2004) noted that trust exists in situations of uncertainty. Boon, Holmes, Hinde and Groebel (1991) wrote that the context for trust requires risk, as did Lewicki, McAllister and Bies (1998). Similarly, Griffin (1967) also claimed that trust occurs in risky conditions.

Trust that is defined based on the context requires a situation where risk is present. If it is defined based on the context and the trustor, that trustor must have some lack of control over the outcome of a situation. If it is defined based on the *trustee* and the context, however, then the trustee must necessarily have some control over the outcome and the ability to do harm to the trustor. If trust is defined entirely by the trustee, their benevolence is the main factor. If it is defined by the trustee and the trustor, then the trustor's confidence and belief in that benevolence is of importance. See Figure 2 for a model of the overlapping and distinct requirements of trust, for each type of definition.

These three main definitional groupings of trust (those referencing the trustor, the trustee, and the shared contextual situation) closely mirror the antecedents of trust (See Hancock et al., 2011, and Schaefer et al., 2016) which are also grouped according to factors related to the trustor, trustee, and context. It is possible, then, that contextual antecedents of trust are reliable predictors of trust *when trust is defined by the context*, and that trustee-related antecedents of trust can be used to anticipate trust levels *when trust is defined based on the trustee*. If this theory is accurate, then the way in which trust is measured must be influenced by the definition being used.

Combining Aspects of Trust

Figure 2 shows the aspects of trust definitions that are important when trust is defined based on the trustee, the trustor, or the context, and any pair of the two. For instance, in order to define trust based on the context and the trustor, the situation must be one in which the trustor, specifically, has some lack of control over the outcome. When trust is based on the *trustee* and the context, the trustee must have some ability to control the outcome. In order to combine all three aspects of trust into one definition, it is necessary to factor in the requirements for all three aspects, as well as for the combinations between them. We therefore conclude that trust exists when an individual is in a situation in which they are vulnerable to the risk of harm due to the actions of another, and yet has confidence that such harm will not be enacted based on their belief in the other actor’s benevolence.

Methods of Trust Measurement

Trust can be measured via survey or behavioral methods. However, whether those methods are valid depends upon the definition of trust being used. When reliance, for instance, is used as a proxy measure for trust, whether it is coded as a trusting behavior may depend on which aspect of trust is being examined. For instance, one’s behavior is integral in some definitions of trust, but that same behavior may not be seen as trust in a situation without risk. Similarly, if one characterizes another person as benevolent, then they would be considered to have some level of trust in that person according to those who define trust by trustee states, but not by those who define trust by the contextual situation. These differences in trust definitions effect the measurement of trust and can have serious consequences when comparing trust levels between different studies.

Summary and Conclusions

Trust has often been thought of as a singular conceptual entity, and it is true that the many definitions of trust have large overlap. However, there are three distinct categories of trust definitions and they differ sufficiently that there are potential ramifications when comparing any measurements taken using different definitions as their basis. For this reason, it is difficult, if not impossible, to compare directly between any two studies examining trust unless those studies both utilize the same measurement and the same definition: Few do. We suggest that, in order to derive any meaningful information from a study of trust, researchers should adopt their definition based on the perspective of their study. We conclude also that it is necessary to derive agreement on one specific definition of trust towards which this work is aimed.

Table 1: Definitions of Trust

“An individual may be said to have trust in the occurrence of an event if he expects its occurrence and his expectation leads to behavior which he perceives to have greater negative motivational consequences if the expectation is not confirmed than positive motivational consequences if it is confirmed.” (Deutsch, 1958; p. 266).
“An expectancy held by an individual or group that the word, promise, verbal or written statement of another individual or group can be relied on.” (Rotter, 1967).
"Reliance upon the characteristics of an object, or the occurrence of an event, or the behaviour of a person in order to achieve a desired but uncertain objective in a risky situation" (Griffin, 1967; p. 105).
“Trust exists to the extent that a person believes another person (or persons) to be benevolent and honest” (Larzelere & Huston, 1980; p. 596).
“A particular level of the subjective probability with which an agent assesses that another agent or group of agents will perform a particular action, both before he can monitor such action (or independently of his capacity ever to be able to monitor it) and in a context in which it affects his own action” (Gambetta, 1988).
“A state involving confident predictions about another’s motives with respect to oneself in situations entailing risk.” (Boon & Holmes1991)
“Interpersonal trust can be defined as the extent to which a person is confident in, and willing to act on the basis of, the words, actions, and decisions of another individual, given faith that the individual is competent and will not act without duly taking into consideration the interests of the truster” (Mcallister, 1995; p. 6)
“The willingness of a party to be vulnerable to the actions of another party based on the expectation that the other will perform a particular action important to the trustor, irrespective of the ability to monitor or control that other party” (Mayer, Davis, & Schoorman, 1995)
“An individual’s belief or a common belief among a group of individuals that another individual or group (a) makes good

faith efforts to behave in accordance with any commitments both explicit and implicit, (b) is honest in whatever negotiations preceded such commitments, and (c) does not take excessive advantage of another even when the opportunity is available" (Cummings & Bromiley, 1996; p. 303)
"One party's willingness to be vulnerable to another party based on the belief that the latter party is (a) competent, (b) open, (c) concerned, and (d) reliable" (Mishra, 1996; p. 265)
"A willingness to increase your vulnerability with another person, whose behavior you cannot control, in a situation in which your potential benefits are much less than your potential losses if the other person abuses your vulnerability." (Zand, 1997; p. 230)
"A psychological state comprising the intention to accept vulnerability based upon positive expectations of the intentions or behavior of another" (Rosseau et al., 1998)
"Confident positive expectations regarding another's conduct in situations of risk" (Lewicki, McAllister, & Bies, 1998; p. 439)
"The expectation that an actor (1) can be relied on to fulfill obligations, (2) will behave in a predictable manner, and (3) will act and negotiate fairly when the possibility for opportunism is present" (Zaheer, McEvily, & Perrone, 1998; p. 143).
"An individual's or group's willingness to be vulnerable to another party based on the confidence that the latter party is benevolent, reliable, competent, honest, and open" (Hoy & Tschannen-Moran, 1999; p.189).
"the attitude that an agent will help achieve an individual's goals in a situation characterized by uncertainty and vulnerability" (Lee & See, 2004).
"degree to which a user or other stakeholder has confidence that a product or system will behave as intended" (ISO/IEC, 2011)
"as the reliance by an agent that actions prejudicial to their well-being will not be undertaken by influential others." (Hancock, Billings, & Schaefer, 2011; p. 24)

REFERENCES

Boon, S.H., & Holmes, B.J. (1991): The dynamics of interpersonal trust: Resolving uncertainty in the face of risk. In RA Hinde and J Groebel (eds.), *Cooperation and Prosocial Behavior*, (190-211). Cambridge, UK: Cambridge University Press.

Cummings, L.L., & Bromiley, P. (1996). The organizational trust inventory (OTI). *Trust in Organizations: Frontiers of Theory and Research*, 302 (330), 39-52.

Deutsch, M. (1958). Trust and suspicion. *Journal of Conflict Resolution*, 2 (4), 265-279.

Gambetta, D. (1988). *Trust: Making and breaking cooperative relations*. Oxford, England: Basil Blackwell.

Giffin, K. (1967). The contribution of studies of source credibility to a theory of interpersonal trust in the

communication process. *Psychological Bulletin*, 68 (2), 104-120.

Hancock, P.A., Billings, D.R., & Schaefer, K.E. (2011). Can you trust your robot?. *Ergonomics in Design*, 19 (3), 24-29.

Hancock, P.A., Billings, D.R., Schaefer, K.E., Chen, J.Y., De Visser, E.J., & Parasuraman, R. (2011). A meta-analysis of factors affecting trust in human-robot interaction. *Human Factors*, 53 (5), 517-527.

Hoy, W.K., & Tschannen-Moran, M. (1999). Five faces of trust: An empirical confirmation in urban elementary schools. *Journal of School Leadership*, 9 (3), 184-208.

ISO/IEC. (2011). ISO/IEC 25010: 2011 Systems and software engineering--Systems and software Quality Requirements and Evaluation (SQuaRE)--System and software quality models.

Larzelere, R.E., & Huston, T.L. (1980). The dyadic trust scale: Toward understanding interpersonal trust in close relationships. *Journal of Marriage and the Family*, 42 (3), 595-604.

Lee, J.D., & See, K.A. (2004). Trust in automation: Designing for appropriate reliance. *Human Factors*, 46 (1), 50-80.

Lewicki, R.J., McAllister, D.J., & Bies, R.J. (1998). Trust and distrust: New relationships and realities. *Academy of Management Review*, 23 (3), 438-458.

Mayer, R.C., Davis, J.H., & Schoorman, F.D. (1995). An integrative model of organizational trust. *Academy of Management Review*, 20 (3), 709-734.

McAllister, D.J. (1995). Affect-and cognition-based trust as foundations for interpersonal cooperation in organizations. *Academy of Management Journal*, 38 (1), 24-59.

Mishra, A.K. 1996. Organizational responses to crisis. The centrality of trust. In: R.M. Kramer and T.M. Tyler (Eds.), *Trust in Organizations*: 261-287. Thousand Oaks, CA: Sage.

Rotter, J.B. (1971). Generalized expectancies for interpersonal trust. *American Psychologist*, 26 (5), 443-452.

Rousseau, D.M., Sitkin, S.B., Burt, R.S., & Camerer, C. (1998). Not so different after all: A cross-discipline view of trust. *Academy of Management Review*, 23 (3), 393-404.

Schaefer, K.E., Chen, J.Y., 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.

Zaheer, A., McEvily, B., & Perrone, V. (1998). Does trust matter? Exploring the effects of interorganizational and interpersonal trust on

performance. *Organization Science*, 9 (2), 141-159.

Zand, D. E., (1997). *The leadership triad: Knowledge, trust, and power*. Oxford University Press on Demand.