

Human Factors: The Journal of the Human Factors and Ergonomics Society

<http://hfs.sagepub.com/>

A Meta-Analysis of Factors Affecting Trust in Human-Robot Interaction

Peter A. Hancock, Deborah R. Billings, Kristin E. Schaefer, Jessie Y. C. Chen, Ewart J. de Visser
and Raja Parasuraman

Human Factors: The Journal of the Human Factors and Ergonomics Society 2011 53: 517

DOI: 10.1177/0018720811417254

The online version of this article can be found at:

<http://hfs.sagepub.com/content/53/5/517>

Published by:



<http://www.sagepublications.com>

On behalf of:



[Human Factors and Ergonomics Society](http://www.hfes.org)

Additional services and information for *Human Factors: The Journal of the Human Factors and Ergonomics Society* can be found at:

Email Alerts: <http://hfs.sagepub.com/cgi/alerts>

Subscriptions: <http://hfs.sagepub.com/subscriptions>

Reprints: <http://www.sagepub.com/journalsReprints.nav>

Permissions: <http://www.sagepub.com/journalsPermissions.nav>

>> [Version of Record](#) - Sep 16, 2011

[What is This?](#)

Report Documentation Page

Form Approved
OMB No. 0704-0188

Public reporting burden for the collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to a penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number.

1. REPORT DATE OCT 2011		2. REPORT TYPE		3. DATES COVERED 00-00-2011 to 00-00-2011	
4. TITLE AND SUBTITLE A Meta-Analysis of Factors Affecting Trust in Human-Robot Interaction				5a. CONTRACT NUMBER	
				5b. GRANT NUMBER	
				5c. PROGRAM ELEMENT NUMBER	
6. AUTHOR(S)				5d. PROJECT NUMBER	
				5e. TASK NUMBER	
				5f. WORK UNIT NUMBER	
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) University of Central Florida, Department of Psychology, Orlando, FL, 32816				8. PERFORMING ORGANIZATION REPORT NUMBER	
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)				10. SPONSOR/MONITOR'S ACRONYM(S)	
				11. SPONSOR/MONITOR'S REPORT NUMBER(S)	
12. DISTRIBUTION/AVAILABILITY STATEMENT Approved for public release; distribution unlimited					
13. SUPPLEMENTARY NOTES					
14. ABSTRACT					
15. SUBJECT TERMS					
16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT unclassified	b. ABSTRACT unclassified	c. THIS PAGE unclassified			

A Meta-Analysis of Factors Affecting Trust in Human-Robot Interaction

Peter A. Hancock, Deborah R. Billings, and Kristin E. Schaefer, University of Central Florida, Jessie Y. C. Chen, U.S. Army Research Laboratory, and Ewart J. de Visser, and Raja Parasuraman, George Mason University

Objective: We evaluate and quantify the effects of human, robot, and environmental factors on perceived trust in human-robot interaction (HRI).

Background: To date, reviews of trust in HRI have been qualitative or descriptive. Our quantitative review provides a fundamental empirical foundation to advance both theory and practice.

Method: Meta-analytic methods were applied to the available literature on trust and HRI. A total of 29 empirical studies were collected, of which 10 met the selection criteria for correlational analysis and 11 for experimental analysis. These studies provided 69 correlational and 47 experimental effect sizes.

Results: The overall correlational effect size for trust was $r = +0.26$, with an experimental effect size of $\bar{d} = +0.71$. The effects of human, robot, and environmental characteristics were examined with an especial evaluation of the robot dimensions of performance and attribute-based factors. The robot performance and attributes were the largest contributors to the development of trust in HRI. Environmental factors played only a moderate role.

Conclusion: Factors related to the robot itself, specifically, its performance, had the greatest current association with trust, and environmental factors were moderately associated. There was little evidence for effects of human-related factors.

Application: The findings provide quantitative estimates of human, robot, and environmental factors influencing HRI trust. Specifically, the current summary provides effect size estimates that are useful in establishing design and training guidelines with reference to robot-related factors of HRI trust. Furthermore, results indicate that improper trust calibration may be mitigated by the manipulation of robot design. However, many future research needs are identified.

Keywords: trust, trust development, robotics, human-robot team

Address correspondence to D. R. Billings, Department of Psychology, University of Central Florida, Orlando, FL 32816; e-mail: dbillings@knights.ucf.edu.

HUMAN FACTORS

Vol. 53, No. 5, October 2011, pp. 517-527

DOI:10.1177/0018720811417254

INTRODUCTION

Human-Robot Partnerships

Robots are frequently used in environments that are unreachable by or are unsafe for human beings. Robotic operations include, among others, planetary exploration, search and rescue, activities that impose hazardous levels of workload on human operators, and actions requiring complex tactical skills and information integration (Hinds, Roberts, & Jones, 2004; Parasuraman, Cosenzo, & de Visser, 2009). Robotic usage is penetrating into many diverse applicational realms, especially in the advanced surgical areas and as assistive technologies for injured and disabled persons (Guizzo & Goldstein, 2005; Heerink, Krose, Evers, & Wielinga, 2010; Tsui & Yanco, 2007). When they are used in military operations, robots are often currently perceived as tools to be manipulated by humans to accomplish specific discrete functions (Chen, Barnes, & Harper-Sciari, 2010).

Yet, as robot capabilities grow, the possibility arises that they might provide higher level functions as full-fledged team members. For this higher functioning to blossom, however, effective human-robot partnerships will need to be forged to ensure success in dangerous conflict situations, particularly because of the increased stress and cognitive workload demands placed on contemporary warfighters (Hancock & Warm, 1989).

In future military contexts, warfighters are likely to be mandated to interact with a diverse inventory of robots on a regular basis, particularly in dynamic and stressful environments (Chen & Terrence, 2009). Already, robotic systems have demonstrated their usefulness in decision making, communication, enhancement of warfighter situation awareness, combat efficiency, and reducing uncertainty in volatile situations (Adams, Bruyn, Houde, & Angelopoulos, 2003). However, the assumption that introducing robots into human teams will result in better performance, as

compared with when the team or robot operates independently, may not always be justified. Although the addition of robotic systems may lead to improved team capabilities, it may also create difficult challenges that need to be overcome before such hybrid partnerships can work more effectively (Adams et al., 2003).

Research continues to address such challenges as creating and validating metrics for the evaluation of a wide spectrum of human-robot interactions (HRI) issues (Steinfeld et al., 2006); designing human-robot interfaces to facilitate interaction, operator understanding, and situation awareness (Chen et al., 2010; Chen, Haas, & Barnes, 2007; Keyes, Micire, Drury, & Yanco, 2010); translating qualities of good human teammates into features of the robot; and encouraging human trust in robots (Groom & Nass, 2007), which is perhaps foremost among these challenges.

Human-Robot Trust

For a human-robot team to accomplish its goal, humans must trust that a robotic teammate will protect the interests and welfare of every other individual on the team. The level of trust in any robotic partner will be particularly critical in high-risk situations, such as combat missions (Groom & Nass, 2007). Trust is important in these contexts because it directly affects the willingness of people to accept robot-produced information, follow robots' suggestions, and thus benefit from the advantages inherent in robotic systems (Freedy, de Visser, Weltman, & Coeyman, 2007). Trust therefore very much affects the decisions that humans make in uncertain or risky environments (Park, Jenkins, & Jiang, 2008). For example, the less an individual trusts a robot, the sooner he or she will intervene as it progresses toward task completion (de Visser, Parasuraman, Freedy, Freedy, & Weltman, 2006; Steinfeld et al., 2006).

However, some accounts from warfighters in the field demonstrate the ease with which trust can actually develop between robots and humans in stressful operations. In fact, one Explosive Ordnance Disposal unit named its robot Sgt. Talon and gave it promotions and even Purple Hearts for its stellar bomb disposal performance (Garreau, 2007). Other accounts by soldiers, in contrast,

illustrate the difficulties in trusting robots in these situations. For instance, the SWORD (Special Weapons Observation Reconnaissance Detection) system was developed and deployed in Iraq in 2007 to support combat operations (Ogreten, Lackey, & Nicholson, 2010). SWORD, although fully operational, was never used in the field because soldiers did not trust it to function appropriately and safely in dangerous situations because of unexpected movements caused by technological malfunctions (Ogreten et al., 2010).

As illustrated by these respective accounts, varying levels of trust in robots currently exist across the HRI domain. Inappropriate levels of trust may have negative consequences, such as overreliance on and misuse of the system (in cases of extremely high levels of trust) or disuse of the system entirely (in cases of very low levels of trust) (Lee & See, 2004; Parasuraman & Manzey, 2010; Parasuraman & Riley, 1997). Both distrust and overreliance can undermine the value of the HRI system. Trust also influences neglect tolerance, which is defined as the decline in semiautonomous robot performance as human attention is directed to other tasks and/or as the task complexity increases (Goodrich, Crandall, & Stimpson, 2003).

When an individual places a large amount of trust in a robot and does not feel compelled to actively manage it, he or she may ignore the robot for long periods. Consequently, neglect tolerance should be appropriately calibrated to the capabilities of the robot and the level of human-robot trust. Too much neglect can make it difficult for the individual to regain situation awareness after redirecting attention back toward the robot. Too little neglect means the human operator is not attending to his or her own personal tasks, thus resulting in suboptimal system performance overall.

Current Research

Trust in HRI is very much related to trust in automation in general, which has been studied with respect to its various performance influences (Chen et al., 2010; Lee & See, 2004; Parasuraman, Sheridan, & Wickens, 2008; Sheridan, 2002). Robots differ from most other automated systems in that they are mobile, are sometimes built in a fashion that approximates human or animal

form, and are often designed to effect action at a distance. Such differences could suggest that human trust may differ for robots versus other forms of automation, although this difference would need to be demonstrated empirically, and few if any such direct comparisons have been conducted to date.

Alternatively, one can begin with the view that human-robot trust and human-automation trust share similar fundamental characteristics but allow for the possibility for differences as new evidence is obtained. Certainly the literature on human-automation trust provides a fertile ground for understanding a number of factors influencing how humans trust other external agents. The human-robot trust literature is more restricted, but nevertheless, sufficient numbers of empirical studies have been conducted to warrant a meta-analysis to identify the major factors currently involved.

Trust can be dynamically influenced by factors (or antecedents) within the robotic system itself, the surrounding operational environment, and the nature and characteristics of the respective human team members (Park et al., 2008). Each of these factors can play an important role in trust development. To date, reviews of trust in HRI have been qualitative and descriptive, and existing experiments largely attempt to extrapolate the optimum degree of trust for a given outcome (e.g., team performance, reliance on the robot). In doing so, a degree of inappropriate trust (i.e., excessive trust or too little trust) is also identified for each potential outcome of HRI, such as over- or underreliance and poor team collaboration. The factors that affect the process by which trust develops in any HRI situation also need to be considered. However, to date, the existing body of knowledge has mostly looked at the momentary state of trust and not its process of development per se. This latter evolution clearly awaits further investigation.

Given the foregoing observations, the goal of the current research was to perform a comprehensive objective and quantitative review of identified antecedents of trust in human-robot teams. Meta-analytic methods were applied to the extant literature on trust and HRI with the aim of quantifying the effects of differing dimensions on human-robot trust. Determining their

relative impact on trust will not only provide an indication of current trends in the human-robot trust research, but will also lead to the identification of areas critical for future study. Consequently, our quantitative review contributes an empirical foundation upon which to advance both theory and practice.

ANALYTICAL METHOD

Sample of Studies

A literature search was conducted using library databases (including PsycINFO, PsycARTICLES, PsycBOOKS, ACM Digital Library, Applied Science and Technology, IEEE, ScienceDirect, and ProQuest Dissertations and Theses). U.S. Army Research Laboratory technical reports were also examined for relevance. In addition, we used a number of web-based search engines, for example, Google and its derivative Google Scholar, to seek further references not discovered by the initial formal scan. The primary search terms included *human-robot interaction*, *robot*, and *trust*. After the initial listing of articles was obtained, reference lists were checked to determine whether any other related studies could be included. In a concurrent process, subject matter experts (SMEs) were consulted for reference to articles that had not been identified by the prior, formal search procedure. SMEs were drawn from military, industry, and academia on the basis of their willingness to participate and their availability.

Following this initial procedure, we examined the collected literature and identified potential factors associated with the development of trust. SMEs also provided guidance in identifying factors influencing trust in human-robot relationships. On the basis of these identified factors, we conducted specific searches in the aforementioned databases using the primary search terms *robot* and *trust* combined with these secondary terms: *prior experience*, *attentional capacity*, *expertise*, *competency*, *personality*, *attitudes*, *propensity to trust*, *self-confidence*, *false alarm*, *failure rate*, *automation*, *anthropomorphism*, *predictability*, *proximity*, *robot personality*, *multi-tasking*, *workload*, *task load*, *culture*, *shared mental models*, and *situation awareness*. When these elicitation processes no longer yielded

new citations, we compiled the final listing of articles. This process resulted in 29 empirical articles, reports, dissertations, and conference proceedings published between 1996 and 2010. Of these, 10 papers containing 69 correlational effect sizes and 11 papers containing 47 experimental effect sizes met selection criteria for inclusion.

Criteria for Study Inclusion

All studies were inspected to ensure that they fulfilled the following four criteria for inclusion in the meta-analysis: (a) Each study had to report an empirical examination of trust in which trust was a directly measured outcome of an experimental manipulation. Studies in which trust served as the experimental manipulation were excluded. (b) The empirical examination of trust was directed toward a robot. Thus, for instance, studies on human-automation trust focusing on a decision aid were excluded because the emphasis of such research is on the decision aid and not a robot, which as discussed earlier can differ in many ways from automated systems in terms of such factors as mobility, sensor and effector capabilities, and so on. (c) The study had to incorporate human participants who either viewed or participated directly in interactions with a robot through physical, virtual, or augmented means. (d) Each study had to include sufficient information to determine effect size estimates.

Papers and literature meeting these criteria are identified in the reference listing in the present article by an asterisk appearing in front of the first author's name (American Psychological Association, 2001). It is important to note that rejecting primary studies in a meta-analysis is a common occurrence and is necessary to ensure meaningful results when combining effect sizes across studies.

Identification of Possible Antecedents of Trust

Studies included in the meta-analysis were classified into three broad categories according to the experimental manipulation: robot-related factors (including performance-based and attribute-based factors), human-related factors (including ability-based and human characteristic factors), and environment-related factors affecting trust

(including team collaboration and task-based factors). These differentiates enabled a quantitative review of the predictive strength of these respective trust factors in human-robot teams. See Figure 1 for factors identified as potential antecedents of human-robot trust on the basis of the literature review and SME guidance.

The Calculation of Effect Size

A meta-analytic approach was used to evaluate the data collected to determine the pattern of findings in the contemporary body of human-robot trust research. First, each study's effect size was calculated using standard formulas (see Hedges & Olkin, 1985; Hunter & Schmidt, 2004; Morris & Deshon, 2002). Studies included in effect size calculation contained both correlational and group design data; therefore the use of multiple meta-analytic methods (correlation and Cohen's d) was necessary. The correlational effects represent an association between trust and the given factor. Cohen's d indicates the standard difference between two means in standard deviation units. From these, we can gather correlational and causal inferences between trust and any given factor. Through both types of meta-analytic effects, the more positive the effect, the more trust. Findings were interpreted with the use of Cohen's (1988) established ranges for small ($d = .20$; $r = .10$), medium ($d = .50$; $r = .25$), and large ($d = .80$; $r = .40$) effect sizes.

Variance Estimates

Several variance estimates were calculated. First, variability of the effect sizes themselves (s_g^2) and variability attributable to sampling error (s_e^2) were estimated. Next, these two values were used to compute the residual variance (s^2). A large (s^2) is an indication that the effect sizes may be heterogeneous and therefore one or more variables are likely to be moderating the magnitude of that particular effect. A final check for homogeneity of variance (s_e^2/s_g^2) was calculated (proportion of total variance accounted for by sampling error). Hunter and Schmidt (2004) suggest that an outcome of 0.75 or greater suggests that the remaining variance is attributable to a variable that could not be controlled for and represents homogeneity of variance. However, large residual variance and small homogeneity

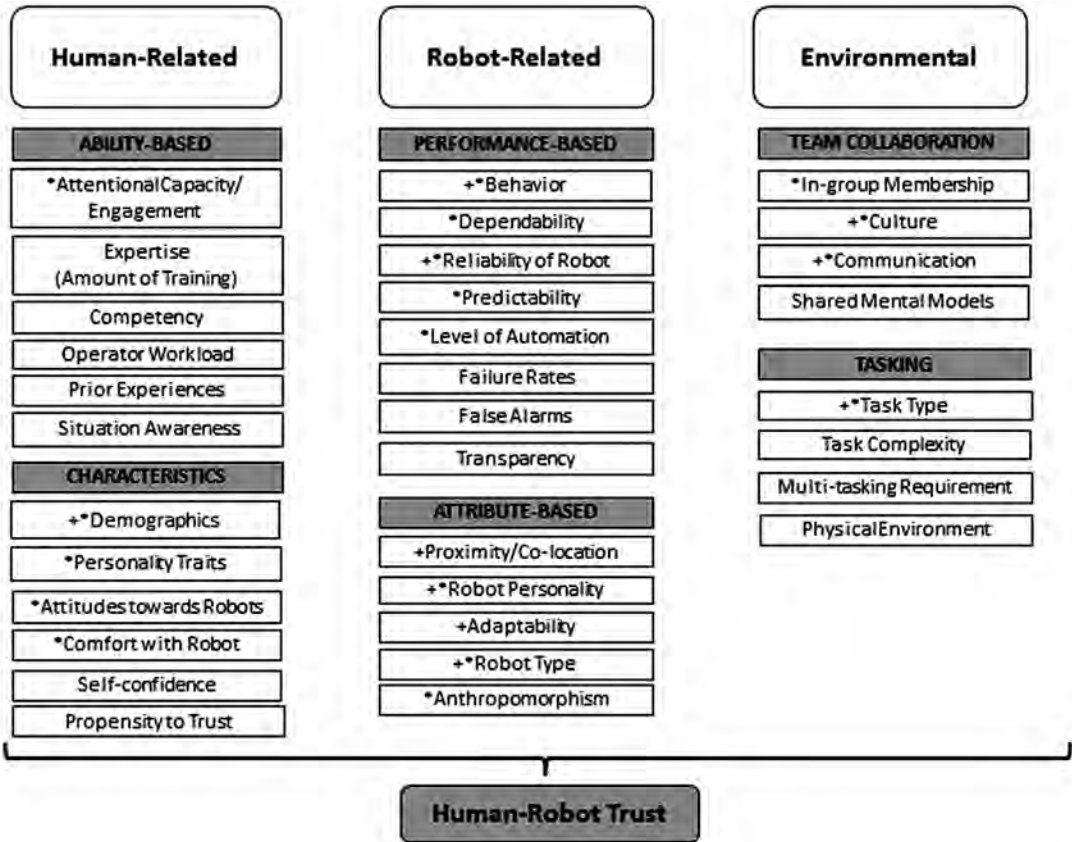


Figure 1. Factors of trust development in human-robot interaction. These factors were identified a priori via literature review and subject matter expert guidance. Factors included in the correlational analysis are starred (*). Factors included in the experimental analysis are crossed (+).

of variance may be seen because of a small number of sample studies, as is evident in some of the following results (see Lipsey & Wilson, 2001, for an in-depth examination of the various strengths and weaknesses relating to meta-analytic procedures).

RESULTS

Overall Outcome Effects

Correlational analysis. For the 10 studies reporting correlational data, the present meta-analytic results indicated that there was a moderate global effect between trust and all factors influencing HRI ($\bar{r} = +0.26$; see Table 1). That the identified confidence interval does not include zero confirms that this identified relationship is consistent and substantive. The

subsidiary analysis between trust and human, robot, and environmental factors individually indicated only small effects for the human dimensions ($\bar{r} = +0.09$) and also the environmental characteristics ($\bar{r} = +0.11$), and because the confidence intervals for human and environmental factors included zero, our current state of knowledge suggests that the human and the environment are not strongly associated with trust development in HRI at this point in time. We should, however, emphasize that these results derive from only a limited number of studies and thus may change with future evaluations.

Robot-related characteristics were found to be moderately associated with trust in HRI ($\bar{r} = +0.24$), in line with the level of the global effect. Robot influences were able to be parsed into two

TABLE 1: Formal Human-Robot Trust Meta-Analysis Results With Correlational Data: Global, Trust Factors, and Robot Factors

Category	<i>k</i>	<i>r̄</i>	s_r^2	s_e^2	s_p^2	s_e^2/s_p^2	95% CI	<i>n</i>
Global	10	+0.26	.14	.01	.13	.05	+0.21 < δ < +0.31	1,228
Trust factors								
Robot	8	+0.24	.21	.01	.20	.05	+0.16 < δ < +0.31	882
Human	7	+0.09	.14	.02	.13	.11	.00 < δ < +0.19	727
Environment	4	+0.11	.11	.01	.10	.08	+0.02 < δ < +0.20	645
Robot factors								
Attribute	5	+0.03	.08	.02	.07	.22	-0.09 < δ < +0.15	686
Performance	5	+0.34	.43	.01	.42	.03	+0.25 < δ < +0.43	607

Note. *k* = number of studies; *n* = sample size; s_r^2 estimates the variability of the effect sizes themselves; s_e^2 estimates the variability attributable to sampling error; s_p^2 is an estimate of the residual variance; s_e^2/s_p^2 is a calculation of homogeneity of variance; CI = confidence interval.

subcategories: robot performance-based factors (e.g., reliability, false alarm rate, failure rate) and attribute-based factors (e.g., proximity, robot personality, and anthropomorphism). With respect to the influence of the robot, it was determined that performance factors were more strongly associated ($\bar{r} = +0.34$) with trust development and maintenance. However, in contrast, robot attributes had only a relatively small associated role ($\bar{r} = +0.03$). Such potential influences for human- and environment-related factors were not examined, as there were insufficient samples at the present juncture to run acceptable quantitative meta-analytic procedures.

Experimental analysis. In contrast to studies that presented correlational data, experimental studies reported group differences. Therefore, for this latter meta-analysis, we used Cohen's *d*. The results for the meta-analytic approach with the use of Cohen's *d* produced a similar pattern to that for the correlational studies. These results, shown in Table 2, indicated there was a large global effect concerning trust and HRI ($\bar{d} = +0.71$). As the confidence interval excluded zero, we can assume this is a substantive and consistently large effect. The subdivision of this global effect into robot, human, and environmental characteristics indicated that the robot ($\bar{d} = +0.67$) had the greatest effect. There was a moderate effect for environmental factors ($\bar{d} = +0.47$) but only very small effects for human factors ($\bar{d} = -0.02$). Robot factors were again parsed into the two

categories of attributes and performance. Robot performance factors ($\bar{d} = +0.71$) were the largest identifiable influences on HRI trust, whereas robot attributes ($\bar{d} = +0.47$) had a smaller but still sizeable influence on trust development.

We should, however, point out that the performance factors are based on two studies, which may bring into question the stability of the effect. However, each study has a sizable effect supporting this found effect. The attribute factors are based on eight studies, pointing to stronger stability of the effect. Specific influential effects for human- and environment-related factors were not examined, as there were insufficient data to run the meta-analysis. In all of the aforementioned categories in which there were sufficient data to identify effects, none of the confidence intervals for the experimental work included zero. Therefore, we can have a degree of confidence that these are each consistent and real effects.

DISCUSSION

Trust is a crucial dimension in maintaining effective relationships with robots. The presence, growth, erosion, and extinction of trust have powerful and lasting effects on how each member of any shared relationship behaves and will behave in the future. Currently, we see technology (and the present panoply of robots) as largely insensate and without individual motive

TABLE 2: Formal Human-Robot Trust Meta-Analysis Results With Cohen's *d*: Global, Trust Factors, and Robot Factors

Category	<i>k</i>	\bar{d}	s_g^2	s_e^2	$s^2\delta$	$s_e^2/s_g^2\delta$	95% CI	<i>n</i>
Global	11	+0.71	.26	.09	.16	.36	+0.53 < δ < +0.89	1,567
Trust factors								
Robot	8	+0.67	.15	.07	.08	.48	+0.48 < δ < +0.85	1,119
Human	2	-0.02	g = +0.01 (Kidd, 2003); g = -0.88 (Scopelliti, Giuliani, & Fornana, 2005)		202			
Environment	5	+0.47	.21	.07	.13	.36	+0.23 < δ < +0.71	609
Robot factors								
Attribute	8	+0.47	.25	.07	.19	.27	+0.28 < δ < +0.65	1,119
Performance	2	+0.71	g = +0.71 (Ross, 2008); g = +0.74 (Tsui, Desai, & Yanco, 2010)		554			

Note. *k* = number of studies; *n* = sample size; s_g^2 estimates the variability of the effect sizes themselves; s_e^2 estimates the variability attributable to sampling error; $s^2\delta$ is an estimate of the residual variance; $s_e^2/s_g^2\delta$ is a calculation of homogeneity of variance; CI = confidence interval.

force. Although we are often frustrated with technological shortcomings and failures and express our frustration accordingly, at heart, we know we are dealing with the residual effects of a remote human designer. However, we stand on the verge of a sufficiently impactful change that our attribution of intentionality to all technology will soon be justified (and see Epley, Waytz, & Cacioppo, 2007; Moravec, 1988). At this juncture, the issue of trust in technological systems will be as influential on social development as it is in our own human-human relationships.

Trust is only one of a number of critical elements essential to human-robot collaboration, but it continues to be a growing concern as robots advance in their functionality. This is especially the case in military and emergency contexts in which a warfighter's or an operator's own life and the lives and safety of others depend on successful interaction. The present research represents one of the first systematic efforts to quantify effects concerning human trust in robots. Our results reveal that robot characteristics, and in particular, performance-based factors, are the largest current influence on perceived trust in HRI. These findings imply that manipulating different aspects of the robot's performance influences trust the most. This finding is central to the consideration of coming robot design. Trends in the literature

indicate that higher trust is associated with higher reliability (for example, see Ross, 2008). Furthermore, the type, size, proximity, and behavior of the robot also affect trust (for examples, see Bainbridge, Hart, Kim, & Scassellati, 2008; Tsui, Desai, & Yanco, 2010). Taking such factors into consideration can have meaningful influence on future robot design and associated human-robot interaction, although further research is still needed to develop specific design heuristics.

Environmental factors were also found to be moderately influential on trust development. Team collaboration characteristics and tasking factors, as identified by SMEs and in the literature itself, were included in this analysis. However, further specification of team-related and task-related effects could not be drawn because of the insufficient number of empirically codable studies. Limited evidence for human-related factors was found. The present findings, however, should not be taken to imply that human characteristics in HRI are not necessarily important. Rather, the small number of studies found in this area suggests a strong need for future experimental efforts on human-related, as well as environment-related, factors.

Although human-automation interaction in general has been researched in more depth (Dzindolet, Peterson, Pomranky, Pierce, & Beck, 2003; Lee & See, 2004; Madhavan & Wiegmann,

2007; Sheridan, 2002; Sheridan & Parasuraman, 2006), sparse empirical research has been conducted in a number of specific and important areas associated with human-robot trust. For instance, as noted, there is a dearth of studies on the human-related characteristics, including prior level of operational experience, attentional capability, the amount of training received, self-confidence, the propensity to trust, existing attitudes toward robots, personality traits, operator workload, situation awareness, and other individual difference factors (see Hancock, Hancock, & Warm, 2009). Gaps in the understanding of the various environmental characteristics include culture (of the team, the individual, and the ambient environment), shared mental models, multi-tasking requirements, task complexity, and task type. We also have limited empirical evidence on the effects of robot “false” alarms. Resolution in these areas is crucial to provide an increasing depth of understanding on trust in HRI.

Our meta-analytic findings have implications for both research and practice. In terms of research, as we build functional models of HRI, we will need to understand and quantify the various influences and derive information on factors we have shown that, to date, are completely missing. Without a larger and active empirical attack, our knowledge will remain precarious and based often on either anecdotal or engineering-centered case studies. With regard to practical implications, the major lesson learned is that a robot’s performance and attributes should be considered the primary drivers of trust. Understanding exactly how these factors affect the development of trust will be critical for trust calibration. For example, we are aware of, and have cited one instance of, a number of occasions in the military in which robots have looked to be deployed, but because of the intrinsic trust question, they have never been taken “out of the box” (often because of a bad reputation preceding the system or its perceived complexity of operation). Consequently, if the perceived risk of using the robot exceeds its perceived benefit, practical operators almost always eschew its use. Hence, training that encourages trust in specific robots is necessary from the point of design inception on, until its eventual field use.

The implications of this research can also be applied to trust in HRI in a number of critical areas outside of the military. These include, especially, considerations in the medical and health care arenas. Assistive robotic technologies are also being developed and tested for mobility purposes, rehabilitation, and of course, surgical aids (Tsui & Yanco, 2007). Social robots for domestic use are also being designed to help people with cognitive and physical challenges maintain a high quality of life (Heerink, et al., 2010). For example, the Care-O-Bot II is a robotic home assistant that was created to support the needs of elderly persons in private homes, allowing individuals greater independence without the necessary presence of a human caregiver (Graf, Hans, & Schraft, 2004).

Other robotic devices, such as wheelchairs (Yanco, 2001) and exoskeletons (e.g., robotic arms, legs; Guizzo & Goldstein, 2005) can benefit disabled individuals by enhancing their remaining physical capabilities. Furthermore, robots are capable of telepresence in life-threatening situations when medical professionals cannot be physically present as well as in cases involving high risk, such as search-and-rescue operations (Murphy, Riddle, & Rasmussen, 2004) and repairs in contaminated conditions, as the recent Japanese nuclear power plant experience has shown (Glionna & Nagano, 2011). Given our current meta-analytic findings, robot performance factors in these domains are also therefore most critical to consider. Regardless of context, a user must trust the robot to enable effective interaction.

Future Research

The type of trust measure used is relevant to the present conclusions. Our meta-analysis found that current trust in HRI is derived almost exclusively via subjective response, measured one time after a specific interaction. However, physiological indicators, such as oxytocin-related measures, and objective measures, such as trust games that assess actual investment behavior, are used frequently in the human-interpersonal trust literature (for examples, see Chang, Doll, van ’t Wout, Frank, & Sanfey, 2010; Keri, Kiss, & Kelemen, 2009). These measures should be

explored in the context of human-robot trust to augment the present perceptual assessments and identify potential inconsistencies between such measures. Discrepancies between an individual's self-report (i.e., perception) and his or her behavior (i.e., observable reaction) is an issue that has been a topic of concern in psychology, and especially applied psychology, for a number of decades (see Hancock, 1996; Natsoulas, 1967). An individual can report that he (or she) will trust a robot, but existing research leads us to believe that this statement-action relationship is not always perfect (Chen & Terrence, 2009). Therefore, empirical research that includes both subjective and objective measurements can provide a more complete portraiture of the genesis and persistence of trust.

A comparison of perceptions and actual robot capabilities is also needed. Each person in a team can have differing perceptions of the intent, performance, and actions of a robotic entity, but indeed, these perceptions may not all match the true capabilities of the robot. These differences in perception may be mitigated to an extent by employing training methods that adequately prepare an individual for the coming interaction. In summary, numerous avenues of research need to be pursued to fully comprehend the role that trust plays in HRI as well as the factors that influence trust in robots themselves as stand-alone entities. Even so, our current findings indicate that currently, the most important element of trust is robot related. Fortunately, these factors (e.g., robot performance, robot attributes) can be directly manipulated by designers (with the constraints of technological capabilities). In this way, we are able to predict to some degree the development of trust in human-robot partnerships in existing systems.

ACKNOWLEDGMENTS

The research reported in this document was performed in connection with Contract No. W911NF-10-2-0016 with the U.S. Army Research Laboratory, under UCF Task No. 3, P. A. Hancock, Principal Investigator. The views and conclusions contained in this document are those of the authors and should not be interpreted as presenting the official policies or position, either expressed or implied, of the U.S. Army Research Laboratory or the U.S. government unless so designated by other authorized

documents. Citation of manufacturer's or trade names does not constitute an official endorsement or approval of the use thereof. The U.S. government is authorized to reproduce and distribute reprints for government purposes notwithstanding any copyright notation herein. We wish to thank the associate editor and two anonymous reviewers for their most helpful comments in revising the present work.

KEY POINTS

- The meta-analytic procedures included 10 references for correlational analysis (yielding 69 correlational effect sizes) and 11 for experimental analysis (yielding 47 experimental effect sizes).
- Robot characteristics, and in particular, robot performance, were found to be the most important influences of trust development.
- Environmental factors moderately influenced trust, whereas little evidence was found for the effect of human characteristics on trust in human-robot interaction (HRI).
- Although human dimensions played a small role in trust development, the lack of findings may be attributable to insufficient empirical data (suggesting limitations in the research), which should be addressed in future experimentation.
- The current summary of findings emphasizes the importance of focusing on robot-related factors in design and training guidelines for HRI.

REFERENCES

- References marked with an asterisk indicate studies included in the meta-analysis.
- Adams, B. D., Bruyn, L. E., Houde, S., & Angelopoulos, P. (2003). *Trust in automated systems literature review* (DRDC Toronto No. CR-2003-096). Toronto, Canada: Defence Research and Development Canada.
- American Psychological Association. (2001). *Publication manual of the American Psychological Association* (5th ed.). Washington, DC: Author.
- *Bainbridge, W. A., Hart, J., Kim, E. S., & Scassellati, B. (2008). The effect of presence on human-robot interaction. In *Proceedings of the 17th IEEE Symposium on Robot and Human Interactive Community* (pp. 701–706). Munich, Germany: IEEE.
- *Biros, D. P., Daly, M., & Gunsch, G. (2004). The influence of task load and automation trust on deception detection. *Group Decision and Negotiation*, 13, 173–189.
- Chang, L. J., Doll, B. B., van 't Wout, M., Frank, M. J., & Sanfey, A. G. (2010). Seeing is believing: Trustworthiness as a dynamic belief. *Cognitive Psychology*, 61, 87–105.
- Chen, J. Y. C., Barnes, M. J., & Harper-Sciarini, M. (2010). Supervisory control of multiple robots: Human-performance issues and user-interface design. *IEEE Transactions on Systems, Man, and Cybernetics—Part C: Applications and Reviews*, 41, 435–454.

- Chen, J. Y. C., Haas, E. C., & Barnes, M. J. (2007). Human performance issues and user interface design for teleoperated robots. *IEEE Transactions on Systems, Man, and Cybernetics-Part C: Applications and Reviews*, 37, 1231-1245.
- Chen, J. Y. C., & Terrence, P. I. (2009). Effects of imperfect automation and individual differences on concurrent performance of military and robotics tasks in a simulated multitasking environment. *Ergonomics*, 52, 907-920.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum.
- *De Ruyter, B., Saini, P., Markopoulos, P., & van Breeman, A. (2005). Assessing the effects of building social intelligence in robotic interface in the home. *Interacting With Computers*, 17, 522-541.
- de Visser, E. J., Parasuraman, R., Freedy, A., Freedy, E., & Weltman, G. (2006). A comprehensive methodology for assessing human-robot team performance for use in training and simulation. In *Proceedings of the 50th Annual Meeting of the Human Factors and Ergonomics Society* (pp. 2639-2643). Santa Monica, CA: Human Factors and Ergonomics Society.
- Dzindolet, M. T., Peterson, S. A., Pomranky, R. A., Pierce, L. G., & Beck, H. P. (2003). The role of trust in automation reliance. *International Journal of Human-Computer Studies*, 58, 697-718.
- Epley, N., Waytz, A., & Cacioppo, J. T. (2007). On seeing human: A three-factor theory of anthropomorphism. *Psychological Review*, 114, 864-886.
- *Evers, V., Maldonado, H., Brodecki, T., & Hinds, P. (2008). Relational vs. group self-construal: Untangling the role of national culture in HRI. In *Proceedings in the 3rd ACM/IEEE International Conference on Human Robot Interaction* (pp. 255-262). New York, NY: Association for Computing Machinery.
- Freedy, A., de Visser, E., Weltman, G., & Coeyman, N. (2007). Measurement of trust in human-robot collaboration. In *Proceedings of the 2007 International Conference on Collaborative Technologies and Systems* (pp. 106-114). Orlando, FL: IEEE.
- Garreau, J. (2007, May 6). Bots on the ground: In the field of battle (or even above it), robots are a soldier's best friend. *The Washington Post*. Retrieved from <http://www.washingtonpost.com>
- Glionna, J. M., & Nagano, Y. (2011, May 8). Japan looks abroad for high-tech help at Fukushima plant. *Los Angeles Times*. Retrieved from <http://articles.latimes.com/2011/may/08/world/la-fg-japan-robots-20110508>
- Goodrich, M. A., Crandall, J. W., & Stimpson, J. (2003, March). *Neglect tolerant teaming: Issues and dilemmas*. Paper presented at the 2003 AAAI Spring Symposium on Human Interaction with Autonomous Systems in Complex Environments, Palo Alto, CA.
- Graf, B., Hans, M., & Schraft, R. D. (2004). Care-O-Bot II development of a next generation robotic home assistant. *Autonomous Robots*, 16, 193-205.
- Groom, V., & Nass, C. (2007). Can robots be teammates? Benchmarks in human-robot teams. *Interaction Studies*, 8, 483-500.
- Guizzo, E., & Goldstein, H. (2005). The rise of the body bots [robotic exoskeletons]. *IEEE Spectrum*, 42(10), 50-56.
- Hancock, P. A. (1996). Effects of control order, augmented feedback, input device and practice on tracking performance and perceived workload. *Ergonomics*, 39, 1146-1162.
- Hancock, P. A., Hancock, G. M., & Warm, J. S. (2009). Individuation: The $N = 1$ revolution. *Theoretical Issues in Ergonomic Science*, 10, 481-488.
- Hancock, P. A., & Warm, J. S. (1989). A dynamic model of stress and sustained attention. *Human Factors*, 31, 519-537.
- Hedges, L. V., & Olkin, I. (1985). *Statistical methods for meta-analysis*. Orlando, FL: Academic Press.
- *Heerink, M., Krose, B., Evers, V., & Wielinga, B. (2010). Assessing acceptance of assistive social agent technology by older adults: The Almere model. *International Journal of Social Robots*, 2, 361-375.
- Hinds, P. J., Roberts, T. L., & Jones, H. (2004). Whose job is it anyway? A study of human-robot interaction in a collaborative task. *Human-Computer Interaction*, 19, 151-181.
- Hunter, J. E., & Schmidt, F. L. (2004). *Methods of meta-analysis: Correcting for error and bias in research findings*. Thousand Oaks, CA: Sage.
- Keri, S., Kiss, I., & Kelemen, O. (2009). Sharing secrets: Oxytocin and trust in schizophrenia. *Social Neuroscience*, 4, 287-293.
- Keyes, B., Micire, M., Drury, J. L., & Yanco, H. A. (2010). Improving human-robot interaction through interface evolution. In D. Chugo (Ed.), *Human-robot interaction* (pp. 183-202). Rijeka, Croatia: InTech.
- *Kidd, C. D. (2003). *Sociable robots: The role of presence and task in human-robot interaction*. (Master's thesis). Massachusetts Institute of Technology, Cambridge.
- *Kidd, C. D., & Breazeal, C. (2004). Effect of a robot on user perception. In *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems* (pp. 3559-3564). Sendai, Japan: IEEE.
- *Kiesler, S., Powers, A., Fussell, S. R., & Torry, C. (2008). Anthropomorphic interactions with a robot and robot-like agents. *Social Cognition*, 26, 169-181.
- Lee, J. D., & See, K. A. (2004). Trust in automation: Designing for appropriate reliance. *Human Factors*, 46, 50-80.
- *Li, D., Rau, P., & Li, Y. (2010). A cross-cultural study: Effect of robot appearance and task. *International Journal of Social Robots*, 2, 175-186.
- Lipsley, M. W., & Wilson, D. B. (2001). *Practical meta-analysis*. Thousand Oaks, CA: Sage.
- *Looije, R., Neerinx, M. A., & Cnossen, F. (2010). Persuasive robotic assistant for health self-management of older adults: Design and evaluation of social behaviors. *International Journal of Human-Computer Studies*, 68, 386-397.
- Madhavan, P., & Wiegmann, D. A. (2007). Similarities and differences between human-human and human-automation trust: An integrative review. *Theoretical Issues in Ergonomics Science*, 8, 277-301.
- Moravec, H. (1988). *Mind children: The future of robot and human intelligence*. Cambridge, MA: Harvard University Press.
- Morris, S. B., & DeShon, R. P. (2002). Combining effect size estimates in meta-analysis with repeated measures and independent groups designs. *Psychological Methods*, 7, 105-125.
- Murphy, R. R., Riddle, D., & Rasmussen, E. (2004). Robot-assisted medical reachback: A survey of how medical personnel expect to interact with rescue robots. In *Proceedings of the 13th Annual IEEE International Workshop on Robot and Human Interactive Communication (RO-MAN 2004)*, (pp. 301-306). Kurashiki, Okayama, Japan: IEEE.
- *Mutlu, B., Yamaoka, F., Kanda, T., Ishiguro, H., & Hagita, N. (2009). Nonverbal leakage in robots: Communication of intentions through seemingly unintentional behavior. In *Proceedings of the 4th ACM/IEEE International Conference on Human Robot Interaction* (pp. 69-76). New York, NY: Association for Computing Machinery.
- Natsoulas, T. (1967). What are perceptual reports about? *Psychological Bulletin*, 67, 249-272.
- Ogreten, S., Lackey, S., & Nicholson, D. (2010, May). *Recommended roles for uninhabited team members within*

- mixed-initiative combat teams*. Paper presented at the 2010 International Symposium on Collaborative Technology Systems, Chicago, IL.
- Parasuraman, R., Cosenzo, K. A., & de Visser, E. (2009). Adaptive automation for human supervision of multiple uninhabited vehicles: Effects on change detection, situation awareness, and mental workload. *Military Psychology, 21*, 270–297.
- Parasuraman, R., & Manzey, D. (2010). Complacency and bias in human use of automation: An attentional integration. *Human Factors, 52*, 381–410.
- Parasuraman, R., & Riley, V. (1997). Humans and automation: Use, misuse, disuse, abuse. *Human Factors, 39*, 230–253.
- Parasuraman, R., Sheridan, T., & Wickens, C. (2008). Situation awareness, mental workload, and trust in automation: Viable, empirically supported cognitive engineering constructs. *Journal of Cognitive Engineering and Decision Making, 2*, 140–160.
- Park, E., Jenkins, Q., & Jiang, X. (2008 September). *Measuring trust of human operators in new generation rescue robots*. Paper presented at the 7th JFPS International Symposium on Fluid Power, Toyama, Japan.
- *Powers, A., Kiesler, S., Fussell, S., & Torrey, C. (2007). Comparing a computer agent with a humanoid robot. In *Proceedings of the International Conference on Human Robot Interaction* (pp. 145–152). Arlington, VA: ACM.
- *Rau, P. L., Li, Y., & Li, D. (2009). Effects of communication style and culture on ability to accept recommendations from robots. *Computers in Human Behavior, 25*, 587–595.
- *Ross, J. M. (2008). *Moderators of trust and reliance across multiple decision aids* (Doctoral dissertation). University of Central Florida, Orlando.
- *Scopelliti, M., Giuliani, M. V., & Fornara, F. (2005). Robots in a domestic setting: A psychological approach. *Universal Access in the Information Society, 4*, 146–155.
- Sheridan, T. B. (2002). *Humans and automation: System design and research issues*. Santa Monica, CA: Wiley.
- Sheridan, T. B., & Parasuraman, R. (2006). Human-automation interaction. *Reviews of Human Factors and Ergonomics, 1*, 89–129.
- Steinfeld, A., Fong, T., Kaber, D., Lewis, M., Scholtz, J., Schultz, A., & Goodrich, M. (2006). Common metrics for human-robot interaction. In *Proceedings of 2006 ACM Conference on Human-Robot Interaction*, (pp. 33–40). Salt Lake City, UT: ACM.
- *Tenney, Y. J., Rogers, W. H., & Pew, R. W. (1998). Pilot opinions of cockpit automation issues. *International Journal of Aviation Psychology, 8*, 103–120.
- *Tsui, K. M., Desai, M., & Yanco, H. A. (2010). Considering the bystander's perspective for indirect human-robot interaction. In *Proceedings of the 5th ACM/IEEE International Conference on Human Robot Interaction* (pp. 129–130). New York, NY: Association for Computing Machinery.
- Tsui, K. M., & Yanco, H. A. (2007). Assistive, surgical, and rehabilitation robots from the perspective of medical and healthcare professionals. In *Proceedings of the AAAI Workshop on Human Implications of Human-Robot Interaction* (pp. 34–39). Vancouver, Canada: AAAI Press.
- *Wang, L., Rau, P. L., Evers, V., Robinson, B. K., & Hinds, P. (2010). When in Rome: The role of culture and context in adherence to robot recommendations. In *Proceedings of the 5th ACM/IEEE International Conference on Human Robot Interaction* (pp. 359–366). New York, NY: Association for Computing Machinery.
- Yanco, H. A. (2001). Development and testing of a robotic wheelchair system for outdoor navigation. In *Proceedings of the 2001 Conference of the Rehabilitation Engineering and Assistive Technology Society of North America*. Arlington, VA: RESNA Press.
- Peter A. Hancock is Provost Distinguished Research Professor and Pegasus Professor in the Department of Psychology and the Institute for Simulation and Training at the University of Central Florida. He received a PhD in motor performance from the University of Illinois, Champaign, in 1983. He is a fellow of and a past president of the Human Factors and Ergonomics Society.
- Deborah R. Billings is a postdoctoral researcher with the Institute for Simulation and Training at the University of Central Florida. She received a PhD in applied experimental and human factors psychology from the University of Central Florida in 2010. She also obtained an MS in modeling and simulation from the University of Central Florida.
- Kristin E. Schaefer is a doctoral student in the modeling and simulation PhD program at the University of Central Florida. She received an MS in modeling and simulation from the University of Central Florida in 2009 and a BA in the field of psychology with a minor in sociology from Susquehanna University.
- Jessie Y. C. Chen is a research psychologist with the U.S. Army Research Laboratory–Human Research and Engineering Directorate (Field Element in Orlando, Florida). She received a PhD in applied experimental and human factors psychology from the University of Central Florida in 2000.
- Ewart J. de Visser is pursuing a PhD degree in human factors and applied cognition at George Mason University. He received his BA in film studies from the University of North Carolina at Wilmington and an MA in human factors and applied cognition from George Mason University in 2007.
- Raja Parasuraman is a university professor in the Department of Psychology at George Mason University and director of the graduate program in human factors and applied cognition. He is director of the Center of Excellence in Neuroergonomics, Technology, and Cognition and chair of the Neuroimaging Core of the Krasnow Institute. He received a PhD in psychology from Aston University, Birmingham, United Kingdom, in 1976.

Date received: February 16, 2011

Date accepted: June 10, 2011