

A Personable Robot: Meta-analysis of Robot Personality and Human Acceptance

Connor Esterwood¹, Kyle Essenmacher¹, Han Yang¹, Fanpan Zeng¹, and Lionel P. Robert²

Abstract—For robots to be of use to humans, they first must be accepted. One important variable that might impact this acceptance is a robot’s personality. To date, results of studies examining robot personality have produced mixed results. One method of making sense of these results is meta-analysis. Therefore, to examine the potential relationship between robot personality and human acceptance, the authors conducted a review of the human–robot interaction literature and leveraged this review for use in a comprehensive meta-analysis. In doing so, this study contributes to the literature in three ways. First, this study found that a robot’s personality does appear to influence humans’ acceptance of robots. Second, this paper provides an introduction to meta-analysis and detailed methodology that can be applied by other researchers. Third, it identifies gaps within the existing literature and presents opportunities for future research.

I. INTRODUCTION

ROBOTS are increasingly positioned to work collaboratively with humans. This collaboration holds many potential benefits including lower human workload [1], increased safety [2], and improved physical and psychological well-being [3]. To date, however, it is unclear whether humans will accept robots [4]–[6]. A lack of human acceptance would undermine any potential benefits associated with human–robot collaboration. Therefore it is important that robot designers implement methods to encourage humans to accept robots.

Incorporating personality into the design of robots is one way to promote acceptance [7]; yet the literature remains largely mixed about whether a robot’s personality actually promotes acceptance. Personality can be defined as an individual’s “characteristic pattern of behaviors in a broad sense (including thoughts, feelings, and motivation)” [8, Pg.527]. A robot’s personality is based on a human interpretation of the robot’s behavior through the human’s observations, interactions, and expectations of the robot [9], [10]. These interpretations have been shown to decrease and other times to increase the acceptance of robots as measured by a wide range of outcomes [9], [11]–[22]. This makes it problematic to form any high-level conclusions on the impact of a robot’s personality on acceptance.

Manuscript received: February, 24, 2022; Revised May, 06, 2022; Accepted May, 09, 2022.

This paper was recommended for publication by Editor Gentiane Venture upon evaluation of the Associate Editor and Reviewers’ comments.

¹Connor Esterwood, Kyle Essenmacher, Han Yang, and Fanpan Zeng are associated with the School of Information, University of Michigan, Ann Arbor, MI 48105, USA (Connor Esterwood* is the corresponding author, email:cte@umich.edu)

²Lionel P. Robert is associated with the School of Information & Robotics Department, University of Michigan, Ann Arbor, MI 48105, USA

Digital Object Identifier (DOI): see top of this page.

To address this problem, we conducted a meta-analysis on robot personality and human acceptance. In particular, we sought to answer the following question: What is the relationship between a robot’s personality and human acceptance? A meta-analysis is an appropriate method to overcome the limitations of any single study. A meta-analysis overcomes such limitations by quantitatively combining the results of multiple studies in order to build an overall estimation of the relationship between variables (i.e., effect size) across a body of literature [23], [24]. In adopting a meta-analytical approach to answering our research question, we contribute findings that go above and beyond existing works such as narrative reviews [7] on this topic. Narrative reviews struggle to make sense of mixed or conflicting results and are largely limited to qualitative interpretation of studies rather than quantitative analysis of the studies [23, Pg.14]. To date, such an analysis exists for humans’ personality and acceptance [5], but no meta-analysis has been conducted of robots’ personality and acceptance. As a result, the findings of this paper contribute to the literature by providing novel insights into whether robot personality can be used to promote robot acceptance.

In addition we also provide a detailed account of how to conduct a meta-analysis. The motivation behind this is twofold. First, meta-analyses are best implemented with an eye to reproducibility because additional data points will inevitably emerge over time and meta-analyses can be built on by future scholars seeking to incorporate new findings [25]. Second, it is our opinion that the field of human–robot interaction (HRI) is ripe for meta-analyses. The number of HRI studies is increasing and their results at times can conflict with one another, leaving unanswered or poorly answered questions. Therefore, in an effort to assist prospective HRI meta-analysts, we detail not only our methodological approach in general but also the specific decisions and choices made while conducting this meta-analysis and why those decisions were made. As a result, this paper contributes to the literature beyond answering a pressing question in HRI by also acting as an introduction to meta-analysis.

This paper is structured as follows. First, we provide a general summary of our methodology used to identify the papers included in our analysis. Second, we detail the meta-analytical approach used and provide justifications and details regarding the different decisions made during this process. Third, we discuss the results of both our review of the literature and our meta-analysis. Finally, we conclude with a discussion of these results and highlight where additional research is needed.

II. METHODOLOGY

The general progression of a meta-analysis begins with a thorough review of the literature. This is necessary because a meta-analysis is only as reliable as the review that supports it. Below, we provide a step-by-step description of the process used to conduct such robust reviews of the literature.

A. Inclusion/Exclusion Criteria

The first step to any review is defining its scope. To accomplish this we set inclusion and exclusion criteria. Specifically, we used a multi-level approach where our inclusion criteria became progressively stricter as we advanced through the screening process. At the highest level, studies were eligible for inclusion if they were classified as academic works (peer-reviewed publications, theses, dissertations, etc.), written in English, and their titles or abstracts contained the terms “robot” and “personality.” We excluded non-English articles primarily because of the relative dominance of the English language in the databases searched and the lack of a specialist translator on the study team. At the second level, studies were eligible if they met all previous requirements, were empirical in nature and design, focused on embodied physical action robots, and included interactions between at least one human and at least one robot. At the third level, we included studies if they met all previous requirements and directly examined robot personalities. Finally, we included studies at the fourth level if they met all previous requirements and used robot personality as an independent variable in their study designs.

In addition to these inclusion criteria, we established a set of exclusion criteria. These exclusion criteria were applied at all points in the screening process. Specifically, we excluded studies if they focused on embodied virtual action agents (e.g., chat-bots), focused on telepresence robots, manipulated a robot’s personality without examining its impact on human subjects, or focused on the negative attitudes toward robots (NARS) measure as the personality of interest. The exclusion of studies that used the NARS scale was based on this scale’s use as a control variable in many studies [7], [26].

B. Search Process

With the inclusion/exclusion criteria in mind, we then needed to conduct a robust search. This requires establishing inclusive search terms, de-duplicating the results, and screening said results based on our inclusion and exclusion criteria. For this study we searched across multiple search engines and databases with the same search terms. Specifically, we conducted this search in 2019 and relied on Google Scholar, the ACM Digital Library, IEEE Explore, and Scopus.

C. Search Terms

To identify studies that examined robot personality, we used the following search terms: “human,” “robot,” “human–robot interaction,” “HRI,” and “personality.” These terms offered a broad search based on previous ad hoc searches and experience with the databases used in this study. We manually screened the results returned from searches using these terms, based on

our stage one inclusion criteria and general exclusion criteria. This screening took the form of paging through each results page until no relevant studies were returned. On average, each page contained 10–25 results by default. In total, we found 1,819 results across all of our searches before accounting for duplicate entries.

D. De-duplication

Given the breadth of our search and the use of multiple search engines that index from the same databases, duplicate studies were likely. As a result, we used a de-duplication strategy relying on the “publish or perish” application [27] and the revtools package in R [28]. In short, we exported search results in .bib format using native export options where possible (ACM Digital Library, IEEE Explore, and Scopus) and the “publish or perish” application where they were not (Google Scholar). The resulting .bib files were then imported into R and merged into a singular data-frame for processing. Duplicates were identified by title via fuzzy matching and followed up with manual screening. After removing duplicates, we were left with 1,069 total unique studies.

E. Screening Procedure

Screening was conducted on the 1,069 unique studies returned from our search as well as an additional 50 studies identified by cross-referencing existing reviews on personality in HRI (see: [7]). We screened studies first based on title, second on abstract, and third on their full-text content. Title screening relied on our first- and second-stage inclusion criteria and general exclusion criteria. We did this manually in the revtools environment, which presented only the titles of articles, with author names and publication names hidden to reduce bias. Studies were then subjected to two abstract screenings. For the first abstract screening, we used the first- and second-stage inclusion criteria and general exclusion criteria. For the second abstract screening, we used the first-, second-, and third-stage inclusion criteria and general exclusion criteria. Studies were then screened one final time, during which we examined their full-text content. For this screening, we used our fourth-stage inclusion criteria as well as all previous stages and our general exclusion criteria, resulting in a total of 13 studies. A summary of this screening procedure with counts at each stage is presented in figure 1 and the studies are displayed in table I.

III. META-ANALYTICAL APPROACH

With the literature review conducted, we then considered meta-analyzing the data from these papers. In particular, for this study we used correlations and adopted a psychometric meta-analytical approach. This involves correcting the distributions of observed correlation coefficients to estimate the distribution of population correlation coefficients. This analysis was conducted in the R environment and relied on the psychmeta package [29]. Using this package, we extracted effect size data and calculated the overall effect sizes weighted by n . In addition, we conducted corrections to these values

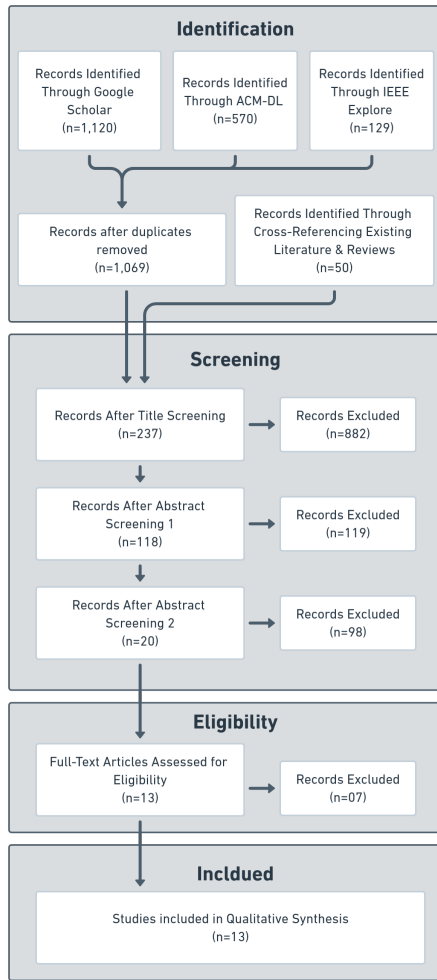


Fig. 1. PRISMA diagram illustrates the screening and study counts.

based on measurement error according to best practices [23]. Where studies reported more than one effect size per sample, we collapsed these studies and took an average [30]. Details of our meta-analytic approach as well as the identification of personality traits and acceptance outcomes are detailed next.

Abbreviation	Where Included	Citation
Kaniasaru & Steinfeld, 2014	Review	[11]
Moshkina & Arkin, 2005	Review	[12]
Leuwerink, 2012	Review & Meta	[13]
Looije <i>et al.</i> , 2010	Review & Meta	[14]
Ogawa <i>et al.</i> , 2009	Review	[15]
Windhouwer, 2012	Review	[16]
Sundar <i>et al.</i> , 2017	Review & Meta	[17]
Gu <i>et al.</i> , 2015	Review & Meta	[18]
Johal <i>et al.</i> , 2014	Review & Meta	[19]
Powers & Kiesler, 2006	Review & Meta	[20]
Park <i>et al.</i> , 2012	Review & Meta	[21]
Lohse <i>et al.</i> , 2008	Review & Meta	[22]
Tay <i>et al.</i> , 2014	Review & Meta	[9]

TABLE I
LIST OF STUDIES INCLUDED IN THIS REVIEW AND/OR META-ANALYSIS
WITH LINK TO CITATION.

A. Identification of Personality Traits

For our meta-analysis to effectively examine the impact of a robot personality on acceptance, we needed an average effect size of the various measures of robot personality across studies. Results of our review, however, indicated that though various measures were present, there was little variety in terms of the personality traits across the literature and even less that contained sufficient data for inclusion in our meta-analysis. Specifically, only three studies looked at personality traits other than extroversion and dominance and, of those, none reported sufficient data for inclusion in our meta-analysis. This indicates a strong need for additional research because personality traits other than extroversion and dominance could not be incorporated into this meta-analysis even though they were present in the broader literature. For the purposes of this paper, we proceeded with our analysis keeping in mind that only a subset of personality traits (extroversion and dominance) were used in the meta-analysis.

B. Identification of Acceptance Outcomes

The outcome of interest for this meta-analysis was robot acceptance. Because acceptance is complex, we first needed to establish a classification for outcomes related to acceptance so the studies' results could be binned (i.e., coded) appropriately. For this meta-analysis, we adopted the perspective of Heerink *et al.* (2009) [31], who extended the Universal Theory of Acceptance and Use of Technology (UTAUT) model originally proposed by Venkatesh *et al.* (2003) [32]. This model defines 13 sub-components of acceptance: anxiety, attitude, facilitating conditions, intention to use, perceived adaptability, perceived enjoyment, perceived ease of use, perceived sociability, perceived usefulness, social influence, social presence, trust, and use/usage. From this list of sub-components, we classified the various acceptance-related outcomes presented across studies.

C. Calculation of Effect Sizes

Effect sizes reflect the magnitude to a treatment effect or the relative strength of a relationship [23]. In a meta-analysis, effect sizes are the primary data point used to consolidate multiple studies and establish an overall effect size between two variables. For this meta-analysis, we utilized correlations (r) values. We obtained these directly from each study's results where available, and where they were not, we calculated them based on the data reported. The specific calculations depended on what data were reported in each study. Generally, we used the available data across studies to calculate Cohen's D and converted this value to r as recommended by Borenstein *et al.* (2011) [23]. For specific equations and further explanation of effect size calculations, see [23] or similarly [24].

D. Reliability

Given that this meta-analysis was conducted in the psychometric tradition, we also examined measurement error (i.e., reliability). Reliability is the ratio of "true" to total variance present in the measurement of a variable of interest [23]. To account for reliability, we took the reliabilities reported across

studies – typically in the form of α values – and incorporated those data points into our overall analysis via R and psychmeta. Where reliability was not reported, we imputed the data from existing reliability data via the recommendations of the psychmeta package [29].

E. Variance Estimates

Variance estimation is critical to a robust meta-analysis because it allows the meta-analyst to make sense of the pattern of effects observed [23]. It does so by examining the degree of consistency across the effect sizes reported and illustrating to what degree these effect sizes result from random error versus genuine differences among studies. Practically, a large degree of unaccounted variance might be attributed to aspects or characteristics of a given study that function as moderators between the predictor and the outcome [23], [24]. Therefore, accounting for the variance present between effect sizes is useful because it can indicate whether moderators are present.

To measure variance between effect sizes, this study used two measures of heterogeneity. Both relied on the Hunter-Schmidt method of variance calculation [33] using the equations recommended by [34]. These produced both a Q statistic and an I^2 value. For Q statistic P values, significant values of $p < 0.05$ indicated that the variability across effect sizes was “real” and not caused by measurement or random error [24] while for I^2 we used the thresholds provided by [35]. Specifically, they state that an I^2 of 0% to 40% indicates that heterogeneity might not be important; 30% to 60% might represent moderate heterogeneity; 50% to 90% may represent substantial heterogeneity; and 75% to 100%, is considerable heterogeneity [35]. We adopted both approaches given the popularity of Q statistics and the relative strength of I^2 values.

IV. RESULTS OF LITERATURE REVIEW

Literature reviews provide useful information in their own right aside from merely setting the groundwork for a meta-analysis. To highlight this, we summarize the literature identified in our review by first describing the characteristics of these studies and then discussing their general findings. Table I provides a list of these studies, where they are incorporated in this paper, and their abbreviations.

A. Study Characteristics

By reviewing the literature on a given topic we can discover not only what has been found relative to independent and dependent variables (IVs/DVs) but also the characteristics of different studies. Specifically, we can determine what IVs and DVs were examined and how. Determining this sets the stage for future research by identifying gaps from both a methodological and conceptual perspectives. To this end, the following sections report sampling information and discuss the IVs and DVs examined across the literature. We detail the gaps that emerged from this examination in section VII.

1) *Sample Sizes & Diversity*: The average sample size across studies was 67, with a standard deviation of 58. Within these samples, gender was not reported by two studies ([18], [20]), with one additional study providing only partial data ([13]). Across the studies that did report gender information, the average ratio of men to women was 48%, indicating a relatively high average level of gender diversity. Figure 2 illustrates these findings. For age groups, five studies ([12], [15], [18]–[20]) did not report this information. The remaining eight studies sampled individuals primarily between ages 18 and 44 ([9], [11], [13], [16], [21], [22]), with only one study examining 45- to 64-year-olds ([14]) and one studying people ages 65+ ([17]). From these results, it appears that the diversity in the ages of samples was relatively low, thus providing an opportunity for additional research.

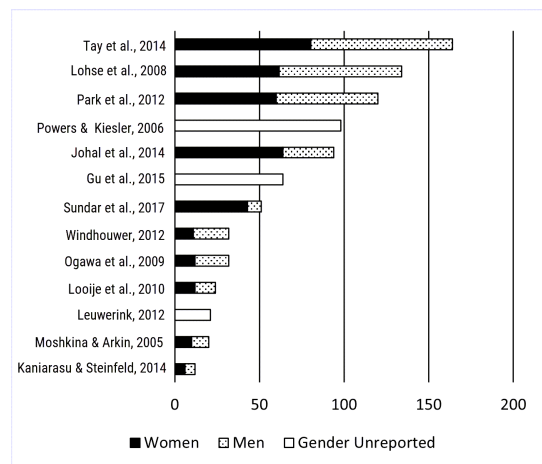


Fig. 2. Bar graph showing sample sizes by study

2) *Personality Traits (IVs)*: The independent variable of interest for this review was robot personality. Results showed that a majority of studies examined personality via the trait of extroversion (k=10) with a minority of studies examining dominance (k=2), openness (k=1), “blame personality” (k=1), and a combination of personality and emotionality (k=1). Measures of robot personality largely relied on custom or new measures of personality (k=4), while established measures such as the Big Five [36], [37] (k=2) and Wiggins scales [38] (k=2) were used less frequently. This finding is similar to that of [10]: extroversion appears to be the most common personality trait examined within the domain of HRI. Table II summarizes these personality traits by study and outcome.

3) *Acceptance (DVs)*: Acceptance was measured via a range of different scales and the majority of studies used more than one measure per sub-component of acceptance. In all, 29 different outcomes related to acceptance were reported. From these, we grouped together outcomes via [31]’s 13 sub-components of acceptance. This produced eight outcomes across these studies. Table II summarizes these outcomes by personality trait and study.

B. Study Findings

Another important contribution of reviews is their capacity for summarizing the existing literature’s results. In doing

so, reviews adopt an interpretivistic approach whereas meta-analyses adopt an empirical approach. Reviews can paint broad pictures of existing literature but often struggle with mixed results, as is the case across the robot personality and acceptance literature. In particular, five studies in our review found that a robot’s personality has a significant impact on acceptance ([11], [12], [14], [18], [19]), two found no significant relationship ([15], [20]), and the remaining six studies found mixed results ([13], [16], [21], [22]) or results that were moderated by a robot’s role ([9], [17]). These results are broken down by sub-component in table II.

Because of these mixed results, this review is somewhat limited. In particular, it can produce a general description of the literature’s findings but it cannot fully determine the presence or relative strength of the relationship between a robot’s personality and human acceptance of the robot. While reporting and summarizing the existing literature are certainly warranted, we now move to our meta-analysis, the larger focus of this paper.

Acceptance Sub-Component	Study	Personality				Sig	Moderators	
		Ext	Dom	Agr	Opn		Moderator	Sig
Acceptance	Looije et al., 2010	x				Y		n/a
	Tay et al., 2014	x				Y	Role	Y
Attitude	Lohse et al., 2008	x				Y		n/a
	Park et al., 2012	x				Y		n/a
	Sundar et al., 2017	x				N		n/a
	Tay et al., 2014	x				Mixed	Role	Y
Enjoyment of Use	Leuwerink, 2012	x				N		n/a
	Lohse et al., 2008	x				N		n/a
	Windhouwer, 2012	x				N.R.	Role	N
	Moshkina & Arkin, 2005					Y		n/a
Ease of Use	Tay et al., 2014	x				N		n/a
	Johal et al., 2014		x			Y		n/a
Usefulness	Leuwerink, 2012	x				Y		n/a
	Lohse et al., 2008	x				Y		n/a
	Sundar et al., 2017	x				N	Role	Y
	Tay et al., 2014	x				Y	Role	Y
	Windhouwer, 2012	x				N.R.	Role	Y
	Tay et al., 2014	x				Mixed	Role	Y
Social Influence	Gu et al., 2015	x				Y		n/a
	Park et al., 2012	x				N		n/a
	Sundar et al., 2017	x				N.R.	Role	Y
Trust	Looije et al., 2010	x				Y		n/a
	Tay et al., 2014	x				Y	Role	Y
	Kaniarasu & Steinfeld, 2014					Y		n/a
Intention to Use	Ogawa_2018	x		x	x	N		n/a
	Park et al., 2012	x				Y		n/a
	Powers_S., 2006		x			N		n/a

TABLE II

ACCEPTANCE OUTCOMES AND PERSONALITY BY STUDY WHERE EXT = EXTROVERSION, DOM = DOMINANCE, AGR = AGREEABLENESS, AND OPN = OPENNESS.

V. RESULTS OF META-ANALYSIS

Leveraging the results of our literature review, we found that of the 13 studies identified, only 9 were eligible for meta-analysis. The other four studies were excluded from the meta-analysis because there was insufficient data reporting such that an effect size could not be calculated ([15]), the effects of robot’s personality could not be isolated ([12], [16]), or the study reported one-off personality traits ([11]). While nine studies is on the lower end of the spectrum for meta-analyses, the number is still sufficient because the canonical minimum number of studies needed for a meta-analysis is only two [39]. Much like sample sizes in experimental studies, however, the power of an analysis can play an important role in determining

the minimum number of studies required for a meta-analysis [40]. Because meta-analysis allows us to critically evaluate and statistically combine the results of these nine studies, we can increase the number of observations and therefore the statistical power present in our analysis. This improves the estimates of the effect size of the treatments. For example, the average power across studies was an α of 0.44, but when these studies were combined into a meta-analysis, the meta-analysis possessed a power of $\alpha = 0.94$. Given that the meta-analysis possessed sufficient power and accurately represented the literature, we considered it sufficient as a standalone meta-analysis for the purposes of this paper.

A. Main Effect of Robot Personality on Acceptance

Table III and figure 3 show the results of a meta-analysis examining the impact of robots’ personalities on acceptance outcomes. In table III, K indicates the number of studies while N represents the number of subjects across studies and the “mean R” indicates the average correlation between robot personality and acceptance. Figure 3 visually represents these results via a forest plot, where the correlation between acceptance and personality is depicted as a point bounded by its confidence interval (lines to the left and right of the point). This was repeated once for each study included in the analysis and can help illustrate whether the effects across studies tend to line up or vary substantially from one study to the next.

Analysis Type	Predictor	Outcome	sig	k	N	Mean R	95% CI
Overall	Robot Personality	Acceptance	Y	9	770	0.23	[0.15,0.39]

TABLE III
RESULTS OF META-ANALYSIS INVESTIGATING THE RELATIONSHIP BETWEEN ROBOT PERSONALITY AND ACCEPTANCE

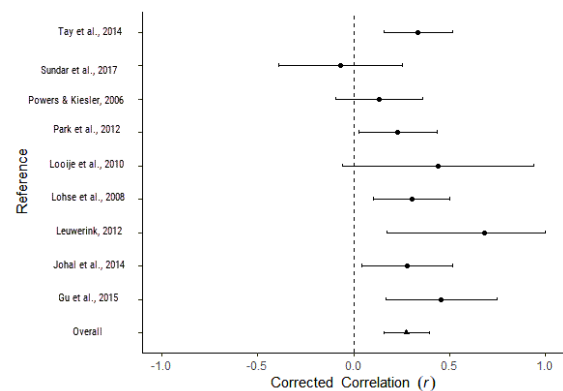


Fig. 3. Forrest plot with effect sizes between robot personality & acceptance.

These results indicate that the estimated corrected score (i.e., “true score”) relationship between robot personality and acceptance is significant and positive ($k=9$, $r^2=0.23$, 95% CI [0.15, 0.39]). Specifically, it appears that a robot’s personality has a significant impact on humans’ acceptance of robots and that this impact might be positive. While this result in and of itself is insightful, it is also useful to check for moderators.

To accomplish this, we conducted an analysis of heterogeneity based on Q and I^2 statistics. Results of this analysis showed that moderators are unlikely. This was the case because we observed a non-significant Q statistic ($Q = 11.78$ $p=0.16$) and a relatively low I^2 value ($I^2 = 32.1$).

B. Robustness Checks

To assess the overall robustness of our meta-analytical results, we investigated the potential for publication bias and assessed the individual impact of any study to skew our results. To do this, we used a funnel plot and also conducted a leave-one-out sensitivity analysis. Next, we explain each approach in detail along with its results.

1) *Publication Bias*: Publication bias is the degree to which “the research that appears in the published literature is systematically unrepresentative of the population of completed studies” [41, Pg.1]. To determine the publication bias present in this meta-analysis, we used funnel plots instead of the fail-safe N or file-drawer analysis. Unlike the funnel plots, the fail-safe or file-drawer analysis methods do not “directly acknowledge the average size of, or the variation in, effects that have been observed” [42, Pg.123]. Further, “different versions of the fail-safe N produce very different results” [42, Pg.124], which frequently leads to misinterpretation and provides only vague indications of publication bias [42].

The funnel plot analysis (in figure 4) shows that there is a greater number of studies with higher correlation (r) values indicating that positive correlations are published more frequently than negative correlations. Each dot in figure 4 represents a single study; the y-axis is the standard error of the estimate while the x-axis is the correlation. Studies with higher power are placed closer to the top and those with lower power are placed toward the bottom of the figure. Investigations of these plots can reveal publication bias when the dots (studies) are grouped together and are found more frequently on the right (positive correlation) side of the straight horizontal line than on the left (negative correlation) side.

Based on this study’s results, it appears that publication bias is present and outlets tend to favor strong and positive correlations. The degree of bias, however, appears to be relatively moderate. This is because some studies appear in the low-to-no-correlation range, indicating that weaker results are also reported. The plot is somewhat asymmetrical, but this is not especially pronounced. Asymmetrical groupings imply the presence of unpublished results, and the stronger the asymmetry, the stronger the publication bias presented. Therefore, it appears that the literature is biased but that the degree of this bias is not extreme.

2) *Leave-One-Out Sensitivity Analysis*: To further assess the robustness of our results, we conducted a leave-one-out sensitivity analysis. This analysis assists in the identification of outliers because we can determine the impact that the exclusion of any single study has on our overall findings. This analysis acts by running nine separate meta-analyses (one per study), excluding a different study each time. Figure 5 presents the results of this analysis, including the mean ρ or correlation coefficient of a meta-analysis conducted with the

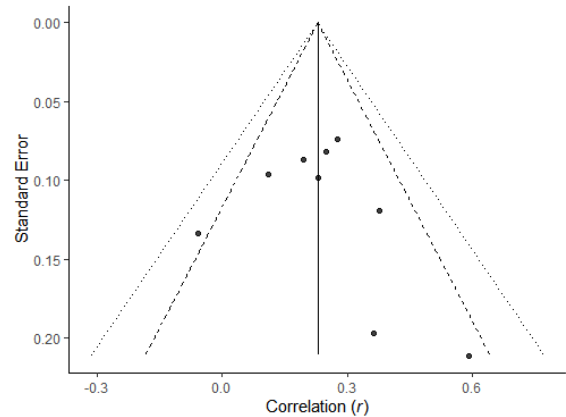


Fig. 4. Funnel plot showing effect sizes and standard errors.

named study left out of the analysis. The point is the mean ρ bounded by 95% and 80% confidence intervals. The point repeats for each study until the analysis has been run once for each excluded study. Inspection of these results shows shifts from the exclusion of different studies, but these shifts are small, indicating that it is unlikely that any individual study produced an out-size effect on our overall findings.

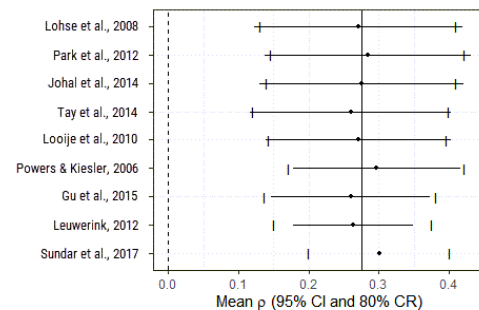


Fig. 5. Leave-one-out analysis comparing impact of each study’s results.

VI. SUMMARY OF FINDINGS

In this literature review and meta-analysis, we investigated the impact of robot personality on acceptance. To do so, we screened more than 1,800 studies, eventually including 13 studies in the qualitative synthesis and nine studies in the quantitative meta-analysis. Results of this meta-analysis indicated that robot personality appears to have a significant and positive impact on acceptance ($k=9$, $r^2=0.23$, 95% CI:[0.15,0.39]).

VII. DISCUSSION & CONTRIBUTIONS

In this paper, we set out with the primary goal of answering the question: *What is the relationship between a robot’s personality and human acceptance?* To accomplish this, we conducted a meta-analysis and detailed our methodology to provide a step-by-step example on how to conduct a meta-analysis. In doing so, we produced two sets of contributions to the literature. The first of these is the results of our meta-analysis itself and the second is the associated methodological outline provided. Next we discuss each of these contributions.

A. Robot Personality and Human Acceptance Insights

The results of the meta-analysis provide evidence that robot personality is significantly and positively related to acceptance but that the effect size is small ($r < 0.30$) [43], [44]. Based on the existing literature, one might have concluded that the relationship was small and negative ($r = -0.1$) [17] or even large and positive ($r = 0.7$) [13]. This would mislead researchers and designers to either abandon or over-rely on robot personality for promoting human acceptance. Results of this study show that robot personality is an important factor to consider but that it would be a mistake to solely rely on it to promote a human's acceptance of robots. For practitioners, this means that when attempting to promote robot acceptance, a robot's personality is one important factor but that other factors should be considered as well.

Aside from this primary finding, our meta-analysis found no evidence of moderators among this set of studies. This was the case because the heterogeneity statistics were non-significant. However, this could be a result of the lack of diversity in the studies' samples. For example, the participants' ages ranged mostly between 18 years and 44 years and did not incorporate minors or older adults (65+). Furthermore, the social context that the human-robot interaction is embedded within has been shown to be an important factor [45]. However, few studies in our meta-analysis clearly defined this social context, nor did these studies examine a wide variety of social contexts. Therefore additional studies with more diverse samples and a wider range of social contexts are needed before we can rule out the possibility of moderators.

Finally, the majority of studies included in our meta-analysis looked at extroversion, with only two examining dominance. Therefore the results of this meta-analysis might speak more to extroversion's impact on acceptance than robot personality generally. It is possible that other personality traits such as neuroticism, openness, agreeableness, or conscientiousness could have a different relationship with acceptance. To determine this, future studies are warranted that examine other personality traits.

B. Meta-Analyses & HRI

To help spur research, in this paper we provide a detailed methodology for meta-analysts in HRI. We do so because the field of HRI has seen a growing influx of research in recent years leading to a great wealth of papers and experimental results. This large sum of literature presents an opportunity for those seeking to make sense of the often disjointed and mixed results found in the HRI literature. It is no surprise, then, that meta-analysis has been used by a handful of scholars to examine topics such as trust [46], [47], anthropomorphism [48], [49], and human personality [5]. There still, however, remain many unanswered questions. As a result, additional contributions to the HRI literature in the form of new meta-analyses on pressing topics such as anthropomorphism, trust repair, individual differences, and more have the potential to produce significant advancements in our understanding of HRI.

C. Limitations & Future Work

While our paper provides a valid approach to meta-analysis, this approach is not without limitations. Specifically, it is possible that additional papers have been published on this topic in 2020 and beyond. These additional papers might not only impact the qualitative synthesis of this literature but also the meta-analytic results. In addition, the studies included in this meta-analysis were written in the English language. Future researchers might wish to investigate non-English-language publications, as well, to determine whether these studies exist and, if so, how their results compare to those found in this paper. Finally, the measures of acceptance and personality used across studies in our meta-analysis varied among studies. While these measures were ultimately similar, a certain degree of variance might be attributable to the differences that do exist. Therefore, future meta-analyses could examine this variance.

REFERENCES

- [1] H. Osawa, A. Ema, H. Hattori, N. Akiya, N. Kanzaki, A. Kubo, T. Koyama, and R. Ichise, "What is real risk and benefit on work with robots? from the analysis of a robot hotel," in *Proceedings of the Companion of the 2017 ACM/IEEE International Conference on human-robot interaction*, 2017, pp. 241–242.
- [2] B. E. Noah and B. N. Walker, "Trust calibration through reliability displays in automated vehicles," in *Proceedings of the Companion of the 2017 ACM/IEEE International Conference on Human-Robot Interaction*, ser. HRI '17. New York, NY, USA: Association for Computing Machinery, 2017, p. 361–362.
- [3] K. S. Welfare, M. R. Hallowell, J. A. Shah, and L. D. Riek, "Consider the human work experience when integrating robotics in the workplace," in *2019 14th ACM/IEEE international conference on human-robot interaction (HRI)*. IEEE, 2019, pp. 75–84.
- [4] S. You and L. Robert, "Emotional attachment, performance, and viability in teams collaborating with embodied physical action (epa) robots," *Journal of the Association for Information Systems*, vol. 19, no. 5, pp. 377–407, 2017.
- [5] C. Esterwood, K. Essenmacher, H. Yang, F. Zeng, and L. Robert, "A meta-analysis of human personality and robot acceptance in human-robot interaction," in *39th ACM Conference on Human Factors in Computing Systems (CHI)*. ACM, 2021.
- [6] C. Esterwood and L. Robert, "Robots and covid-19: Re-imagining human-robot collaborative work in terms of reducing risks to essential workers," *ROBONOMICS: The Journal of the Automated Economy*, vol. 1, pp. 9–9, 2021.
- [7] L. P. Robert, R. Alahmad, C. Esterwood, S. Kim, S. You, and Q. Zhang, "A review of personality in human-robot interactions," *Foundations and Trends® in Information Systems*, vol. 4, no. 2, pp. 107–212, 2020.
- [8] A. Baumert, M. Schmitt, M. Perugini, W. Johnson, G. Blum, P. Borke-nau, G. Costantini, J. J. Denissen, W. Fleeson, B. Grafton *et al.*, "Integrating personality structure, personality process, and personality development," *European Journal of Personality*, vol. 31, no. 5, pp. 503–528, 2017.
- [9] B. Tay, Y. Jung, and T. Park, "When stereotypes meet robots: The double-edge sword of robot gender and personality in human-robot interaction," *Computers in Human Behavior*, vol. 38, pp. 75–84, 2014.
- [10] C. Esterwood and L. P. Robert, "A systematic review of human and robot personality in health care human-robot interaction," *Frontiers in Robotics and AI*, p. 306, 2021.
- [11] P. Kaniarasu and A. M. Steinfeld, "Effects of blame on trust in human robot interaction," in *The 23rd IEEE International Symposium on Robot and Human Interactive Communication*. IEEE, 2014, Conference Proceedings.
- [12] L. Moshkina and R. Arkin, "Human perspective on affective robotic behavior: a longitudinal study," in *2005 IEEE/RSJ International Conference on Intelligent Robots and Systems*. IEEE, 2005, Conference Proceedings.
- [13] K. Leuwerink, "A robot with personality: Interacting with a group of humans," in *Proceedings of the 16th twente student conference on IT*, vol. 4, 2012, Conference Proceedings.

- [14] R. Looije, M. A. Neerinx, and F. Cnossen, "Persuasive robotic assistant for health self-management of older adults: Design and evaluation of social behaviors," *International Journal of Human-Computer Studies*, vol. 68, no. 6, pp. 386–397, 2010.
- [15] K. Ogawa, C. Bartneck, D. Sakamoto, T. Kanda, T. Ono, and H. Ishiguro, *Can an Android Persuade You?* Singapore: Springer Singapore, 2018, book section Chapter 14, pp. 235–247.
- [16] D. Windhouwer, "The effects of the task context on the perceived personality of a nao robot," in *Proceedings of the 16th twente student conference on IT*, 2012, Conference Proceedings, pp. 1–7.
- [17] S. S. Sundar, E. H. Jung, T. F. Waddell, and K. J. Kim, "Cheery companions or serious assistants? role and demeanor congruity as predictors of robot attraction and use intentions among senior citizens," *International Journal of Human-Computer Studies*, vol. 97, pp. 88–97, 2017.
- [18] J. Gu, T. Kim, and Y. Kwon, "Am i have to extrovert personality? an empirical investigation of robot's personality on the two contexts," *Indian Journal of Science and Technology*, vol. 8, no. 26, pp. 1–11, 2015.
- [19] W. Johal, S. Pesty, and G. Calvary, "Towards companion robots behaving with style," in *The 23rd IEEE International Symposium on Robot and Human Interactive Communication*, 2014, Conference Proceedings, pp. 1063–1068.
- [20] A. Powers and S. Kiesler, "The advisor robot: tracing people's mental model from a robot's physical attributes," in *Proceedings of the 1st ACM SIGCHI/SIGART conference on Human-robot interaction*, 2006, Conference Proceedings, pp. 218–225.
- [21] E. Park, D. Jin, and A. P. del Pobil, "The law of attraction in human-robot interaction," *International Journal of Advanced Robotic Systems*, vol. 9, no. 2, p. 35, 2012.
- [22] M. Lohse, M. Hanheide, B. Wrede, M. L. Walters, K. L. Koay, D. S. Syrdal, A. Green, H. Huttenrauch, K. Dautenhahn, G. Sagerer, and K. Severinson-Eklundh, "Evaluating extrovert and introvert behaviour of a domestic robot; a video study," in *RO-MAN 2008 - The 17th IEEE International Symposium on Robot and Human Interactive Communication*. IEEE, 2008, Conference Proceedings.
- [23] M. Borenstein, L. V. Hedges, J. P. Higgins, and H. R. Rothstein, *Introduction to meta-analysis*. John Wiley & Sons, 2011.
- [24] M. W. Lipsey and D. B. Wilson, *Practical meta-analysis*. SAGE publications, Inc, 2001.
- [25] T. D. Pigott and J. R. Polanin, "Methodological guidance paper: High-quality meta-analysis in a systematic review," *Review of Educational Research*, vol. 90, no. 1, pp. 24–46, 2020.
- [26] S. You and L. Robert, "Teaming up with robots: An imoi (inputs-mediators-outputs-inputs) framework of human-robot teamwork," *Int.J Robotic Eng(IJRE)*, vol. 2, no. 3, 2018.
- [27] A. W. Harzing, "Publish or perish app," 2007. [Online]. Available: <https://harzing.com/resources/publish-or-perish>
- [28] M. J. Westgate, "revtools: An r package to support article screening for evidence synthesis," *Research synthesis methods*, 2019.
- [29] J. A. Dahlke and B. M. Wiernik, "psychmeta: An r package for psychometric meta-analysis," *Applied Psychological Measurement*, vol. 43, no. 5, pp. 415–416, 2019.
- [30] J. Higgins and S. Green, "How to include multiple groups from one study," *Cochrane handbook for systematic reviews of interventions version*, vol. 5, no. 0, 2008.
- [31] M. Heerink, B. Krose, V. Evers, and B. Wielinga, "Measuring acceptance of an assistive social robot: a suggested toolkit," in *RO-MAN 2009-The 18th IEEE International Symposium on Robot and Human Interactive Communication*. IEEE, 2009, pp. 528–533.
- [32] V. Venkatesh, M. G. Morris, G. B. Davis, and F. D. Davis, "User acceptance of information technology: Toward a unified view," *MISQ*, pp. 425–478, 2003.
- [33] J. E. Hunter and F. L. Schmidt, *Methods of meta-analysis: Correcting error and bias in research findings*. Sage, 2004.
- [34] J. P. Higgins, S. G. Thompson, J. J. Deeks, and D. G. Altman, "Measuring inconsistency in meta-analyses," *Bmj*, vol. 327, no. 7414, pp. 557–560, 2003.
- [35] J. J. Deeks, J. P. Higgins, and D. G. Altman, "9.5.2 identifying and measuring heterogeneity," *The Cochrane handbook for systematic reviews of interventions. Version*, vol. 5, no. 0, 2011.
- [36] G. Saucier, "Mini-markers: A brief version of goldberg's unipolar big-five markers," *Journal of personality assessment*, vol. 63, no. 3, pp. 506–516, 1994.
- [37] A. Van Vliet, "Effecten van teamsamenstelling op het moreel van uitgezonden militairen (effects of teamcomposition on the morale of patched soldiers)(rep. no. tm-01-a0404)," *Soesterberg: TNO technische menskunde*, 2001.
- [38] J. S. Wiggins, "A psychological taxonomy of trait-descriptive terms: The interpersonal domain," *Journal of personality and social psychology*, vol. 37, no. 3, p. 395, 1979.
- [39] J. C. Valentine, T. D. Pigott, and H. R. Rothstein, "How many studies do you need? a primer on statistical power for meta-analysis," *Journal of Educational and Behavioral Statistics*, vol. 35, no. 2, pp. 215–247, 2010.
- [40] T. D. Pigott, *Introduction*. Boston, MA: Springer US, 2012, pp. 1–6.
- [41] H. R. Rothstein, A. J. Sutton, and M. Borenstein, "Publication bias in meta-analysis," *Publication bias in meta-analysis: Prevention, assessment and adjustments*, pp. 1–7, 2005.
- [42] B. J. Becker, "Failsafe n or file-drawer number," *Publication bias in meta-analysis: Prevention, assessment and adjustments*, pp. 111–125, 2005.
- [43] J. Cohen, "A power primer," *Psychological bulletin*, vol. 112, no. 1, p. 155, 1992.
- [44] J. Kneer, "Effect sizes," in *The SAGE Encyclopedia of Communication Research Methods*, M. Allen, Ed. SAGE Publications, 2017, pp. 407–407.
- [45] M. Jousse, M. Lohse, J. G. Pérez, and V. Evers, "What you do is who you are: The role of task context in perceived social robot personality," in *2013 IEEE International Conference on Robotics and Automation*. IEEE, 2013, pp. 2134–2139.
- [46] P. A. Hancock, D. R. Billings, K. E. Schaefer, J. Y. Chen, E. J. De Visser, and R. Parasuraman, "A meta-analysis of factors affecting trust in human-robot interaction," *Human factors*, vol. 53, no. 5, pp. 517–527, 2011.
- [47] P. Hancock, T. T. Kessler, A. D. Kaplan, J. C. Brill, and J. L. Szalma, "Evolving trust in robots: specification through sequential and comparative meta-analyses," *Human factors*, vol. 63, no. 7, pp. 1196–1229, 2021.
- [48] A. Weiss and C. Bartneck, "Meta analysis of the usage of the godspeed questionnaire series," in *2015 24th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*. IEEE, 2015, pp. 381–388.
- [49] E. Roesler, D. Manzey, and L. Onnasch, "A meta-analysis on the effectiveness of anthropomorphism in human-robot interaction," *Science Robotics*, vol. 6, no. 58, p. eabj5425, 2021.