A Meta-Analysis of Human Personality and Robot Acceptance in Human-Robot Interaction

CONNOR ESTERWOOD, School of Information, University of Michigan, USA

KYLE ESSENMACHER, School of Information, University of Michigan, USA

HAN YANG, School of Information, University of Michigan, USA

FANPAN ZENG, School of Information, University of Michigan, USA

LIONEL P. ROBERT, School of Information, Robotics Institute, University of Michigan, USA

ABSTRACT:

Human personality has been identified as a predictor of robot acceptance in the human–robot interaction (HRI) literature. Despite this, the HRI literature has provided mixed support for this assertion. To better understand the relationship between human personality and robot acceptance, this paper conducts a meta-analysis of 26 studies. Results found a positive relationship between human personality and robot acceptance. However, this relationship varied greatly by the specific personality trait along with the study sample's age, gender diversity, task, and global region. This meta-analysis also identified gaps in the literature. Namely, additional studies are needed that investigate both the big five personality traits and other personality traits, examine a more diverse age range, and utilize samples from previously unexamined regions of the globe.

1 INTRODUCTION

In the study of human-robot interaction, human personality has been identified as a predictor of whether humans will accept robots. Examples include studies that employ a human's personality trait score to predict whether that person will accept a robot in a given context [2, 3, 8, 15, 18, 35, 53, 59, 86, 87, 90, 100, 102]. There is now a tendency to assume that at least one of the Big Five personality traits can be used to understand whether a person will accept a robot [81, 82]. Despite this intuitive appeal, several literature reviews on the topic have highlighted inconsistencies in the relationship between human personality and the acceptance of robots [28, 82]. However, none of those studies engaged in a systematic quantitative meta-analysis to empirically examine this relationship. This leaves open questions regarding

the importance of human personality in predicting the acceptance of robots. In the absence of such an analysis, it is difficult to know whether or when human personality can be used to predict robot acceptance. It is also difficult to confidently deploy effective design solutions based on this knowledge.

The goal of this study was to conduct a quantitative meta-analysis to answer these questions: *Can human personality be used to predict robot acceptance, and if so, when?* This paper focuses on personality and robot acceptance because they represent two vital concepts in our broader nomological networks used to understand human-robot interactions. Human personality is a stable human trait that is always present across different interactions and contexts, supporting its use as a robust predictor of human attitudes and behaviors [11, 29, 32, 62, 89]. Technology acceptance has long been associated with the ability to fully leverage that technology [83]. More specifically, robot acceptance has been positively associated with human–robot interactions and successful human–robot collaborations [13, 44, 67]. We conducted a meta-analysis because it could allow us to answer these questions (i.e. Can human personality be used to predict robot acceptance, and if so, when?) in the face of conflicting or mixed results by leveraging multiple studies on human personality and robot acceptance. This meta-analysis allowed us to overcome the limitations of any single HRI study suffering from small sample and effect sizes (see:[82]). A meta-analysis can also help to identify gaps in the literature that need further study. In all, this meta-analysis allowed us to expand and clarify our understanding of the relationship between two essential concepts used to view human interactions with robots.

2 BACKGROUND

Personality can be defined as an individual's "characteristic pattern of behaviours in the broad sense (including thoughts, feelings, and motivation)" [6, Pg.527]. Therefore, it is not surprising that personality can predict an individual's future cognition, emotion and behavior [57, 75]. Personality has been especially useful when attempting to comprehend the ways in which people respond to others and the quality of these interactions overall [25, 75, 96]. Scholars across several disciplines including sociology [33, 68], psychology [9, 27], political science [23, 34, 91], human computer interaction (HCI) [55, 77, 99], and organizational behavior [5, 27, 75] have each examined the ability of personality to predict certain human behaviors.

Theories of personality can be generally categorized into five schools of thought [66]. Although each school of thought is unique and distinct they also overlap in many ways. To date, there is no single unified theory of personality but, rather a range of approaches with different emphases. A comprehensive review and discussion of each approach is outside the scope of this paper (see [66] for review). However, we present a summary of each of these approaches.

The first approach is the trait based approach to personality psychology. The trait based approach takes the view that sets of traits or "organized dispositions within the individual" [66, Pg.30], make up the backbone of personality and can be utilized to predict human behavior [1, 65]. There is debate within this school of thought on the precise number of traits as well as which sets of traits most accurately represent personality. Different sets of traits each have some form of empirical support but the set garnering the most empirical support is the Big Five [66]. The Big Five personality traits are: extraversion, neuroticism, openness, agreeableness, and conscientiousness [50, 63]. The trait based approach to personality typically relies on self-reports that categorize individuals based on their scores across trait-related items [66]. The second approach is the psycho-dynamic approach to personality psychology. This approach emphasizes the importance of emotional responses and emotional conflict [31]. This approach also emphasizes the role of the unconscious mind [66]. The third approach is the cognitive/social learning approach. This approach takes the view that personality is formed via cognition. The cognitive/social learning approach focuses on individual differences related to thinking, planning, and problem solving [66, 69, 70]. The fourth approach is the humanistic view

of personality. This approach sees personality as a manifestation of searches for meaning, personal constructs, life narrative, and self-actualization [66]. Finally, there is the evolutionary psychology approach. This approach revolves around personality as a neurological or biological mechanism developed for evolutionary purposes. This approach focuses on biological stimuli and responses as a means of explaining personality and individual behavior [14, 66]. Although these schools of thought differ in distinct ways, they all see personality as a predictor of human behavior. Each of these theories differentiates the mechanisms through which personality is formed and expressed, but they hold in common a belief that personality exists, differs among individuals, and can lead to differences in behavior.

For the proposes of this review we adopted the trait based approach to personality. We used this approach because of its popularity in the HRI literature. This perspective views human personality as more than just a single factor and capable of being divided into a variety of different personality traits. This division or specification allows researchers a more precise approach to linking a particular personality trait to a particular outcome [42, 66, 94]. At present, the Big Five personality traits are the predominant set of traits used across many fields [57, 81, 82], including HRI studies [28, 55, 77, 81, 82, 99]. The personality traits that make up the Big Five are extroversion, agreeableness, conscientiousness, neuroticism (sometimes called emotional instability), and openness to experience [36, 50, 65]. Agreeableness can be defined as the extent to which someone is cooperative and friendly [75]. Conscientiousness is the extent that individuals are careful, deliberative, and self-aware of their actions [94]. Extroversion is often presented as a spectrum with two poles one being extroversion and the other introversion. Extroversion is the extent to which an individual is outgoing, assertive, talkative, and sociable [80], whereas introversion is the degree to which someone is shy, enjoys being quiet, and enjoys being alone [25]. Neuroticism is the degree to which someone is easily angered, not well-adjusted, insecure, or lacking in self-confidence [25]. Finally, openness to experience is often defined as the degree to which one is imaginative, curious, and broadminded [64].

Robot acceptance is a common outcome of interest in many human-robot interaction (HRI) personality studies. Acceptance can be conceptualized in a variety of ways. This paper defines robot acceptance as "the demonstrable willingness within a user group to employ technology for the tasks it is designed to support" [24, Pg.1]. Based on this definition, prior work has examined and applied aspects of the commonly employed the Unified Theory of Acceptance and Use of Technology (UTAUT) model [98] to many different domains [26, 52]. Heerink et al. [44], applied the UTAUT model to the HRI domain and provided a "toolkit" with which to examine acceptance of robots. This model's components and a discussion are covered in Section 4.2.

Research on personality and robot acceptance has found mixed results. Of the personality traits investigated, the most measured trait is extroversion. These studies typically found positive correlations between extroversion and acceptance [2, 3, 8, 15, 18, 86, 87], with the exception of two studies that found negative correlations between some acceptance outcomes and extroversion [53, 102], thereby painting a more complex picture of these relationships. In terms of neuroticism, existing studies differed significantly from one another. Most studies looking at this personality trait and acceptance outcomes identified negative correlations between neuroticism and robot acceptance [8, 86, 90, 102] whereas others found positive correlations [53, 87, 100]. The trait of openness also saw contradiction but only from one study indicating negative correlations [53]; the remainder argued for a positive, albeit weak, correlation [8, 12, 18, 86]. The least commonly measured personality traits were agreeableness and conscientiousness. Results on agreeableness were largely consistent showing a positive correlation with acceptance [8, 18, 53, 86, 93]; in constrast, [21] provided evidence of a negative correlation. Results related to conscientiousness were mixed with some showing positive correlations [8, 21, 86] and others showing negative [53, 59]. Taken as a whole, it is clear that there are mixed results on the subject thus warranting further investigation of the matter.

Many researchers have sought to examine these connections. The field of HRI has seen increasing discussion of these variables as visible in the results of [81] and [82]. Across this literature we see connections and compelling links between human personality and acceptance but, the overall mixed nature of these results presents new challenges to interpretation. A systematic quantitative meta-analysis was missing from this literature. Influenced by the recent success of meta-analyses to help draw causal inference [38, 47, 60, 105], we sought to apply this method to the subject at hand. To accomplish this, we used a psychometric meta-analytic approach to investigate the relationship between human personality and acceptance of robots across the existing literature. In this approach, we tackled inconsistency within the literature by weighting studies on the basis of sample size and adjusting for measurement error. Our results contribute to the literature by providing a touchstone from which future researchers can more easily identify gaps and contribute further to the understanding of human personality as it relates to robot acceptance.

3 SELECTION OF CANDIDATE STUDIES

3.1 Search Process

The literature search employed multiple searches via Google Scholar, the ACM Digital Library, IEEE Explore, and Scopus.

3.2 Search Terms

Using nine search terms, we reviewed results on the search engine results page (SERP), where we applied specific inclusion and exclusion criteria. We paged through SERPs progressively until no single result on the list met the specified criteria. Results prior to the page with no relevant results were extracted for review while subsequent results were not. Each SERP contained 10–25 results (depending on the database) by default. In total, we found 6,140 results across all of our searches before accounting for duplicate entries.

The search terms utilized were:

- (human) AND (robot) AND ("personality")
- ("human robot interaction") AND ("personality")
- ("HRI") AND ("personality")
- (human) AND (robot) AND (personality)
- ("human robot interaction") AND (personality)
- ("HRI") AND (personality)
- (robot and personality)
- (human robot interaction and personality)
- (HRI and personality)

These search terms above were selected after a review of the associated keywords from previously selected papers on the topic of personality and HRI. The goal was to identify consistently and frequently utilized keywords. The words "personality", "personalities", "robot" and "human robot interaction" and "HRI" were by far the most consistently and frequently utilized keywords. In addition, we found identical results across search terms for either "human-robot interaction" vs. "human robot interaction". To verify these search results, we compared the articles returned in this search to the articles identified by [82]. Our search terms returned all the articles identified by [82] and 2,254 additional unique articles.

4

3.3 De-Duplication

We exported search results from Google Scholar in .ris format using the "publish or perish" application [41] and imported them into R for processing. For other databases, we exported results using their respective built-in tools. We conducted de-duplication using the revtools package in R [104]. Then, we removed duplicate articles on the basis of title fuzzy matching and followed up with manual screening. Then we removed duplicates, leaving 2,474 unique entries.

3.4 Criteria for Study Inclusion

For this paper we used a three-stage approach, starting with broad eligibility criteria and applying progressively stricter criteria on the previously refined results. We used the initial eligibility criteria in the page-by-page review of search results and in all subsequent screenings. Then we implemented the secondary eligibility criteria in title screening, abstract screening, and full-text screening. The tertiary eligibility criteria were implemented in full-text screening. We used the exclusion criteria throughout all steps of this review.

Papers were initially selected for inclusion if they met three specific criteria. First, studies were required to be classified as articles or academic works. Second, studies were required to be written in the English language. The reason for excluding non-English-language publications relates to the lack of a specialist or translator on the review team, making these studies difficult to screen appropriately. The third criterion for our initial eligibility was that the titles or abstracts retrieved were required to explicitly mention both the term "robot" and the term "personality."

At the secondary level, we selected papers on the basis of four additional eligibility criteria. First, studies were required to be empirical in nature and design. Second, these studies were required to use embodied physical action (EPA) robots. Third, studies were required to include measures of human or robot personality. Fourth, studies were eligible only if they involved humans interacting with the selected EPA robots.

At the tertiary level we used criteria specific to the purposes of our meta-analysis. These eligibility criteria were that studies must examine human personality, investigate its relationship with broadly defined acceptance outcomes, and report data that are useful in the conducting of a meta-analysis (effect sizes, sample size, etc.).

We excluded studies if they (1) focused on embodied virtual action (EVA; i.e., virtual agents), (2) focused on telepresence robots, (3) focused only on manipulating robot personality without examining its impact on humans, or (4) focused only on negative attitudes toward robots (NARS) as the personality trait 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 (see [106], [82]).

3.5 Screening Procedure

We conducted title screening manually in the revtools environment on the 2,474 unique entries previously identified. Screening was done only on the article title with author names and publication name hidden. Title screening was conducted based on the initial eligibility criterion. This screening identified 332 eligible studies.

We conducted abstract screening manually in the revtools environment on the 332 eligible studies. Abstracts were extracted from Google Scholar and manually added to the data-set utilized by revtools. We adopted this approach because Google Scholar has no native export system and the exporting of abstracts on behalf of "publish or perish" is incomplete and contains missing data. This screening incorporated all previous eligibility criteria in addition to the secondary eligibility criteria and led to the identification of 87 relevant studies.

In addition, we identified 50 other potential references from previously published review papers on the topic (see [82]). All papers identified via this means were reviewed in the same way as the papers identified by our search (title and

Paper Abbreviation	Citation Link
Brandstetter_2017	[12]
Chevalier_2015	[15]
Chidambaram_2012	[16]
Conti_ 2017	[18]
Cruz-Maya_2016_Nutrition	[19]
Damholdt_2015	[21]
Gockley_2006	[35]
Haring_2014	[39]
Ivaldi_2017	[49]
Bernotat_2017	[8]
Kimoto_2016	[53]
MacDorman_2015	[61]
Nomura_2007	[72]
Nomura_2008	[71]
Ogawa_2009	[73]
Park_2012	[74]
Reich-Stiebert_2015	[79]
Rosenthal_2013	[84]
Salem_2015	[87]
Sehili_2014	[90]
Syrdal_2006	[92]
Takayama_2009	[93]
Vollmer_2015	[100]
Walters_2005	[101]
Walters_2008	[102]
Weiss_2012	[103]

Table 1. Selected paper for inclusion in meta-analysis with hyperlink to reference

abstract screening) and with identical criteria. Ultimately, 34 of the additional 50 references were eligible for full-text screening.

Full-text screening involved reading each of the 121 selected papers in detail to determine their suitability based on all previously listed criteria and our tertiary eligibility criteria. After completing this screening, 26 papers met all of our eligibility requirements and were included in the meta-analysis. Figure 1 visually represents this review process and the associated counts. Table 1 contains the list of selected papers included in our meta-analysis.

4 ANALYTICAL APPROACH

The meta-analysis in this paper adopted a psychometric approach. This method was utilized to correct for a range of different errors. Specifically, we corrected for three error types: (1) sampling error — random variance from the population value, (2) measure error — inaccuracies that arise from flaws associated with measurement, and (3) reporting error — where a study's reported values differ from its actual observed values. To accomplish this we included, where possible, information about the size of these errors. This information was given for each study (i.e., each correlation) individually (typically via a reliability metric). As a result, each correlation was corrected individually as determined by the R psychmeta package [20] on the basis of the provided reliability information. This correction provided an adjusted
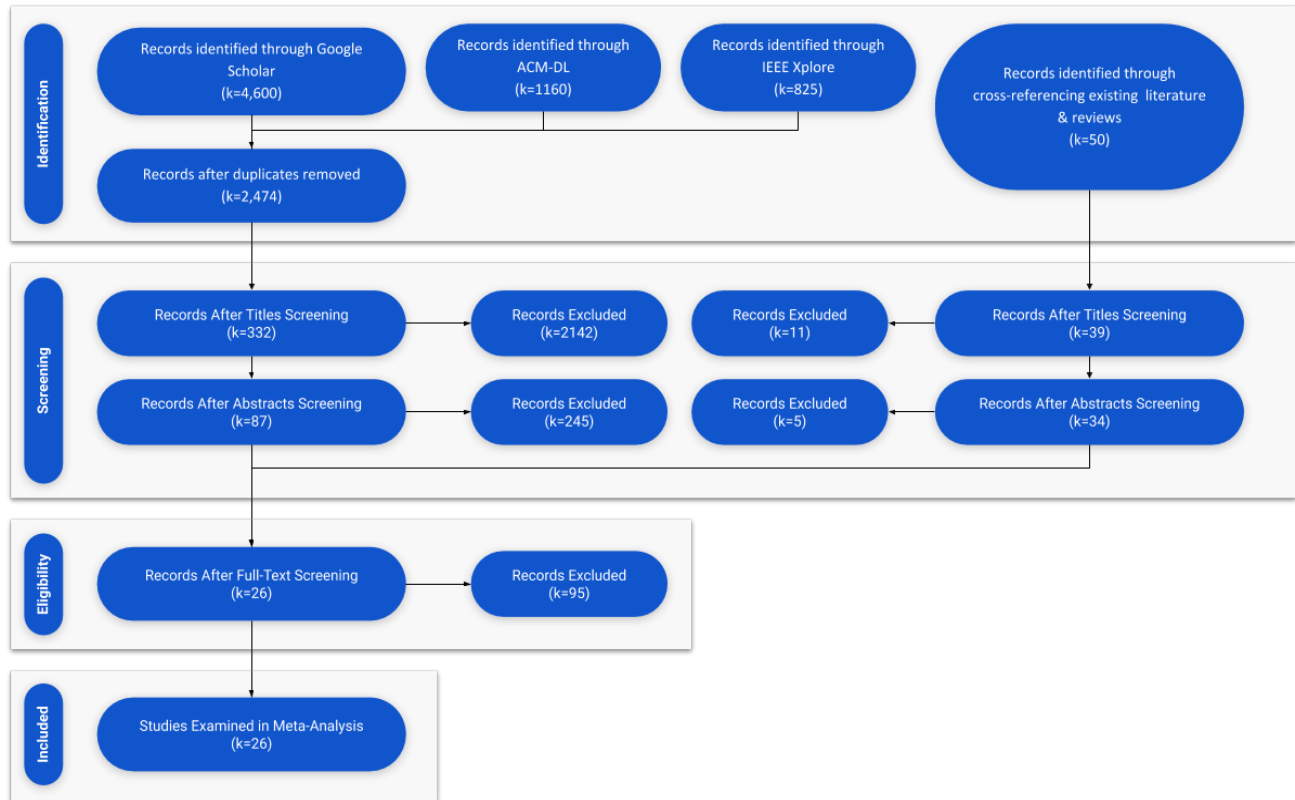
Fig. 1. Prisma Flow Diagram of Literature Selection Process

Code	Construct	Definition	Alpha
ANX	Anxiety	Evoking anxious or emotional reactions when using the system.	.812
ATT	Attitude	Positive or negative feelings about the appliance of the technology.	.851
FC	Facilitating conditions	Objective factors in the environment that facilitate using the system.	.707
TTU	Intention to use	The outspoken intention to use the system over a longer period in time.	.948
PAD	Perceived adaptability	The perceived ability of the system to be adaptive to the changing needs of the user.	.709
PENJ	Perceived enjoyment	Feelings of joy or pleasure associated by the user with the use of the system.	.774
PEOU	Perceived ease of use	The degree to which the user believes that using the system would be free of effort	.726
PS	Perceived sociability	The perceived ability of the system to perform sociable behavior.	.878
PU	Perceived usefulness	The degree to which a person believes that using the system would enhance his or her daily activities	.865
SI	Social influence	The user's perception of how people who are important to him think about him using the system	.794
SP	Social presence	The experience of sensing a social entity when interacting with the system.	.816
Trust	Trust	The belief that the system performs with personal integrity and reliability.	.732
Use	Use/Usage	The actual use of the system over a longer period in time	na

Table 2. Acceptance constructs as presented in [44]

effect size value for each study. We then conducted the meta-analysis on these corrected effect sizes. In cases where studies reported on more than one effect size per sample, we collapsed (reduced down to a single effect size) these studies and took the average [45].

4.1 Identification of Personality Trait

To examine the impact of personality this study used the average effect size of the various measures of human personality across the studies. These personality measures were one of the Big Five personality traits [37, 51]. Most studies reported at least one of these personality traits. We also reverse coded the effects of neuroticism to convert it to emotional stability. This was to ensure that all the effects were compatible with one another. All of this allowed us to investigate the relationship between human personality and robot acceptance across all the studies.

4.2 Identification of Acceptance Outcomes

The primary outcome of interest for this analysis was robot acceptance. For the purpose of this meta-analysis the extended version of the UTAUT model proposed and validated by [44] provided a useful framework by which to define and categorize different subgroups of acceptance. This model defines 13 subgroups of acceptance. The studies identified within this meta-analysis employed a variety of acceptance measures. However, each of these measures was capable of fitting into the subgroups of the UTAUT model provided by [44]. Table 2 details the subgroups, their respective definitions and reported α as presented in [44].

To verify that the theoretical grouping of the outcomes was valid, we conducted an additional analysis. Based on a recommendation by [10], we conducted an analysis of variance (ANOVA) with human personality as the predictor, the overall acceptance construct as the outcome and the outcome type representing each specific outcome under the umbrella of acceptance as a moderator. Significant effects for the moderator would indicate that a specific outcome under the umbrella of acceptance was sufficiently different from the other outcomes. Once identified this outcome should be removed or dropped from the group of outcomes representing acceptance. Non-significant effects would indicate that an outcome under the umbrella of acceptance was sufficiently similar and that it should be included. The results of the aforementioned analysis indicated no significant differences, thereby providing empirical support for the grouping of our selected outcomes under the umbrella of acceptance. Results of this analysis are presented in Table 3. As shown in Table 3, "subgroup 1" and "subgroup 2" are the different groups that fall under the umbrella of acceptance.

Subgroup 1	Subgroup 2	Mean Diff	Sig	CI_LL_95	CI_UL_95
ANX	ATT	0.066226	N	-0.1283	0.2608
ANX	PS	-0.11139	N	-0.3052	0.0824
ANX	PU	-0.06061	N	-0.1952	0.074
ANX	SP	-0.09057	N	-0.3175	0.1363
ANX	Trust	-0.09086	N	-0.2624	0.0807
ANX	TTU	0.162284	N	-0.168	0.4926
ANX	USE	-0.24009	N	-0.6485	0.1684
ATT	PS	-0.17762	N	-0.396	0.0408
ATT	PU	-0.12684	N	-0.2942	0.0406
ATT	SP	-0.15679	N	-0.3772	0.0636
ATT	Trust	-0.15709	N	-0.3572	0.0431
ATT	TTU	0.096058	N	-0.2328	0.4249
ATT	USE	-0.30632	N	-0.6522	0.0395
PS	PU	0.050777	N	-0.1176	0.2192
PS	SP	0.020824	N	-0.2122	0.2538
PS	Trust	0.02053	N	-0.1774	0.2184
PS	TTU	0.273675	N	-0.0635	0.6109
PS	USE	-0.1287	N	-0.508	0.2506
PU	SP	-0.02995	N	-0.3139	0.254
PU	Trust	-0.03025	N	-0.1697	0.1092
PU	TTU	0.222898	N	-0.1071	0.5529
PU	USE	-0.17948	N	-0.7775	0.4185
SP	Trust	-0.00029	N	-0.2195	0.2189
SP	TTU	0.252851	N	-0.1143	0.62
SP	USE	-0.14953	N	-0.8112	0.5121
Trust	TTU	0.253145	N	-0.0755	0.5818
Trust	USE	-0.14923	N	-0.5325	0.2341
TTU	USE	-0.40238	N	-0.8761	0.0713

Table 3. Results of a Wald test comparing correlations for bins with more than k=1 studies

The "Mean Diff" is the difference in values between those groups, while "Sig" indicates if the difference between the two groups are significant, and "CI LL 95" as well as "CI UL 95" represent the 95% confidence interval at the lower and upper levels for that difference, respectively.

4.3 Calculation of Effect Sizes

Effect sizes are defined as values reflective of the magnitude (i.e. strength) of a treatment effect on a relationship between two variables [10]. Effect sizes are a key component to conducting meta-analyses. For this analysis, effect sizes were converted to r (correlation) values because this meta-analysis focuses on correlation. When possible, r was obtained directly from the study. However, when not possible, we calculated it using [17] or [56]. The specific calculations used depended on the data reported. Most approaches involved converting a given statistical result to Cohen's D (non-correlation effect size) and then converting this value to r. We calculated the conversion from Cohen's D to r as recommended by [10] using Equation 1. For conversion to Cohen's D from studies that reported t tests with equal sample sizes, we used Equation 2; for t test values that reported unequal sample sizes, we used Equation 3, and for F test values, Equation 4 per [95], where subscript t refers to treatment and subscript c refers to comparison/control

condition. For studies that reported only η^2 (eta squared) values we used Equation 5 and converted directly to r as recommended by [54]. Eta squared is analogous to r^2 and represents effect size in a similar fashion [88].

$$r = \frac{d}{\sqrt{d^2 + \frac{(n_1 + n_2)^2}{n_1 n_2}}} \tag{1}$$

$$d = \frac{\bar{X}_t - \bar{X}_c}{S_{pooled}} \tag{2}$$

$$d = t\sqrt{\frac{n_t + n_c}{n_t n_c} \frac{n_t + n_c}{n_t + n_c - 2}} \tag{3}$$

$$d = \frac{\bar{X}_t - \bar{X}_c}{\sqrt{MSE(\frac{n_t + n_c - 2}{n_t + n_c})}} \tag{4}$$

$$r = \sqrt{\eta^2} \tag{5}$$

4.4 Reliability

In addition to effect sizes, construct reliability is a vital component to conducting the meta-analysis. Reliability is essentially the ratio of "true" (i.e., corrected) variance to total variance present in the measurement of a variable of interest [10]. When possible, we took construct reliability directly from the study. This typically took the form of αs. When these values were not available, however, we imputed the data from existing reliability data using the recommended method within the R psychmeta package [20].

4.5 Variance Estimates

Another key component to conducting meta-analyses is variance estimates. Variance estimates represent the degree to which the effect sizes observed are caused by random error vs. "real" among between studies [10]. Variance estimates can be used to determine the presence of possible moderators. A large degree of unaccounted variance might not be caused by error alone but instead could be attributable to aspects or characteristics of the study or sample that function as moderators between the predictor and the outcome of interest [10, 58]. Therefore, a significantly large variance estimates indicates possible moderators. To determine whether variance estimates were significantly large, we used two measures of heterogeneity. Both measures relied on the Hunter-Schmidt method of calculating variance [48] and were calculated in R via the psychmeta package [20] using Equation 6 per [43] and Equation 7 per [46]. Details on the specific approach to calculating these values are available via the psychmeta [20] documentation accessible on CRAN. For interpretation we used both Q statistic P values and I^2s. Q statistic, otherwise known as Cochran's Q test, is based on a chi-square distribution which is based on the underlying null hypothesis that the treatment effect size is the same across all the studies [58]. For Q statistic P values, significant values of p < 0.05 indicate that the variability across effect sizes is "real" and not caused by measurement or random error (see Equation 6)[58]. A limitation of Q statistic is that it might be under-powered when few studies are examined in the meta-analysis [58]. The I^2 is a more recent

approach to examining the presence of possible moderators. The I^2 divides the Q statistic minus its degrees of freedom by the Q-value itself (see Equation 7, where I^2 values of 0% to 40% might be considered indicates of low heterogeneity, values between 30% to 60% moderate heterogeneity, 50% to 90% substantial heterogeneity, and 75% to 100% considerable heterogeneity) [22].

$$Q = \sum w_i (ES_i - \bar{ES})^2 \qquad (6)$$

$$I^2 = \left(\frac{Q - df}{Q} \right) * 100 \qquad (7)$$

5 RESULTS

5.1 Main Effect of Personality and Acceptance

Table 4 and Figure 2 present the meta-analytic results for the relationship between the measure of personality effect sizes and robot acceptance. In Table 4, K represents the number of studies included, N represents the number of subjects across all studies "Mean r" is the average correlation value between human personality and acceptance, "CI LL" and "CI UL" are the confidence interval for the lower level (LL) and upper level (UL) for the "mean r" relationship. Figure 2 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 is 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.

Investigating personality and acceptance, the estimated corrected score (i.e., "true score") relationship showed a significant positive relationship between human personality and robot acceptance (k=26, r^2=0.16, 95% CI;[0.09,0.37], where number of studies = k, mean correlation = r^2, and confidence interval = CI). Based on the available data we can conclude that human personality is significantly related to robot acceptance. Specifically, these results indicate a positive correlation between personality and acceptance. Next, we conducted an analysis of heterogeneity based on Q and I^2 to check for possible moderators. The results for the personality measure (Q = 110 p<0.001, I^2 = 76.5) provided compelling evidence for the presence of moderators.

Predictor	Outcome	Sig	k	N	Mean_r	CI_LL_95	CI_UL_95
Personality	Acceptance	Y	26	1611	0.14	0.063	0.3384

Table 4. Results of Primary Meta-Analysis Investigating the Relationship between Human Personality and Acceptance

5.2 Moderators

Given the available data, we examined the role of personality trait type, age, gender, global region, and task type as potential moderators between human personality and robot acceptance. Moderators were coded by multiple researchers, who reached consensus through iterative discussions until all coders agreed on the classification of the moderators. Using this approach, the coders attained close to 100% inter-rater reliability. Among the studies identified, all reported at least one of the Big Five personality traits, 15 reported age, 18 reported gender diversity, 20 reported location or region, and 16 detailed the task type. These are detailed in Table 5.
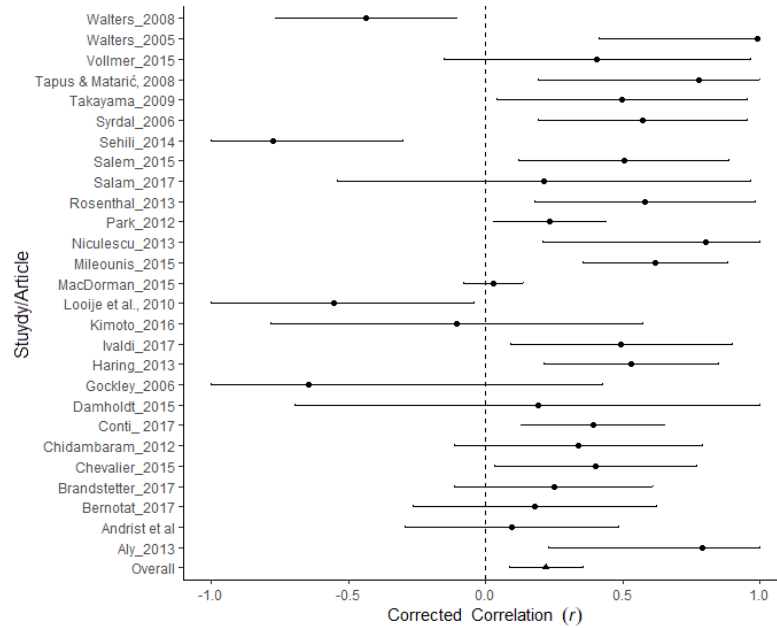
Fig. 2. Forrest plot indicating average effect sizes (r) between personality and acceptance of robots

	Number of Studies	Studies
Personality Trait Type	26	Aly_2013 Andrist et al, Bernotat_2017, MacDorman_2015, Brandstetter_2017, Chevalier_2015, Rosenthal_2013, Syrdal_2006, Conti_ 2017, Gockley_2006, Kimoto_2016, Ivaldi_2017, Haring_2013, Walters_2005, Chidambaram_2012, Niculescu_2013, Park_2012, Damholdt_2015, Takayama_2009, Tapus & Mataric, 2008, Looije et al., 2010, Salam_2017, Salem_2015, Sehili_2014, Vollmer_2015, Walters_2008
Age	15	Andrist et al, Bernotat_2017, Brandstetter_2017, Rosenthal_2013, Syrdal_2006, Conti_ 2017, Kimoto_2016, Ivaldi_2017, Haring_2013, Chidambaram_2012, Takayama_2009, Looije et al., 2010, Salem_2015, Sehili_2014, Walters_2008
Gender	18	Andrist et al, Bernotat_2017, Brandstetter_2017, Chevalier_2015, Rosenthal_2013, Syrdal_2006, Conti_ 2017,Gockley_2006, Kimoto_2016, Ivaldi_2017, Haring_2013, Chidambaram_2012, Park_2012, Takayama_2009, Salem_2015, Sehili_2014, Vollmer_2015, Walters_2008
Global Region	20	Aly_2013, Andrist et al, Bernotat_2017, MacDorman_2015, Brandstetter_2017, Chevalier_2015, Syrdal_2006, Conti_ 2017, Ivaldi_2017, Haring_2013, Walters_2005, Mileounis_2015, Chidambaram_2012, Niculescu_2013, Park_2012, Damholdt_2015, Takayama_2009, Salem_2015, Sehili_2014, Walters_2008
Type of Task	16	Bernotat_2017, Brandstetter_2017, Chevalier_2015, Gockley_2006, Kimoto_2016, Ivaldi_2017, Haring_2013, Mileounis_2015, Chidambaram_2012 Niculescu_2013, Park_2012, Damholdt_2015, Salem_2015, Sehili_2014, Vollmer_2015, Walters_2008

Table 5. Counts and associated studies for each moderator investigated

5.2.1 Personality Trait. Specifically, as seen in Table 6, the "corrected score" relationships among agreeableness (k=6, $\bar{r}^2$=0.20, 95% CI [0.08,0.53]) , extroversion (k=18, $\bar{r}^2$=0.25, 95% CI [0.19,0.53]), and openness (k=8, $\bar{r}^2$=0.24, 95% CI [0.06,0.63]) were all significant and positive, providing evidence of their relationship with robot acceptance. Specifically,

12

these results indicate a positive correlation between agreeableness, extroversion and openness, and robot acceptance. In other words, the more agreeable, extroverted, and open minded an individual was, the more likely he or she was to accept robots. Conscientiousness (k=5, r^2=-0.02, 95% CI [-0.54,0.50]) and neuroticism (k=11, r^2=0.03, 95% CI [-0.15,0.24]) were not significant. In other words, the meta analysis provided no evidence that the more conscientious or neurotic (or emotionally stable) individuals were, the more or less willing they were to accept a robot. In all, the results only found evidence linking three of the Big Five personality traits to robot acceptance. Details of these results are presented in Table 6 and graphically depicted in Figure 3.

Analysis Type	Predictor	Trait Type	sig	k	N	mean_r	CI_LL_95	CI_UL_95
Overall	Personality	All Traits	Y	26	1611	0.14	0.0714	0.3363
Simple Moderator	Personality	Agreeableness	Y	6	243	0.2	0.07	0.4954
Simple Moderator	Personality	Conscientiousness	N	5	123	-0.02	-0.5022	0.4536
Simple Moderator	Personality	Extroversion	Y	18	819	0.25	0.1842	0.477
Simple Moderator	Personality	Neuroticism	N	11	991	0.03	-0.1441	0.2402
Simple Moderator	Personality	Openness	Y	8	322	0.24	0.039	0.6111

Table 6. Results of Personality Trait as Moderator Between Human Personality and Acceptance

To check for possible moderators, we conducted a heterogeneity analysis based on Q and I^2. Results of this test indicated that moderation was likely for conscientiousness (Q = 7.21 p<0.12, I^2 = 44.6), extroversion (Q = 47 p<0.001, I^2 = 64.1), neuroticism (Q = 36.6 p<0.001, I^2 = 72.7), and openness (Q = 16.9 p<0.02, I^2 = 58.5).

5.2.2   Age. To examine the potential moderating impact of age, we grouped studies according to their reported sample's average age and placed them into one of three age groups (18-44, 45-64, 65+). As shown in Figure 4 the relationship between personality and robot acceptance was only significant in the 18-24 age group. Specifically, results indicate that the positive correlation between personality and robot acceptance was only significant for studies with samples of individuals ranging from age 18 and 24. Details of these results are available in Table 7 and Figure 4. It should be noted that a lack of studies prevented us from examining the impact of age by a specific personality trait.

Analysis Type	Predictor	Age Range	Sig	k	N	mean_r	CI_LL_95	CI_UL_95
Overall	Personality	All Ages	Y	15	686	0.16	0.0052	0.4661
Age Group 1	Personality	18-24	Y	4	167	0.27	0.0281	0.6548
Age Group 2	Personality	25-44	N	8	354	0.19	-0.0663	0.6226
Age Group 3	Personality	45-64	N	3	178	0.12	-0.7666	1.0741

Table 7. Results of Age Group as Moderator Between Human Personality and Acceptance ($k \geq 3$)

Analysis Type	Predictor	Gender Diversity	sig	k	N	mean_r	CI_LL_95	CI_UL_95
Overall	Personality	All Levels	Y	18	888	0.19	0.097	0.4463
Gender Div Grp 1	Personality	High	Y	10	465	0.25	0.2454	0.4758
Gender Div Grp 2	Personality	Low	N	4	231	-0.06	-1.0011	0.8269
Gender Div Grp 3	Personality	Medium	Y	4	192	0.34	0.0879	0.7419

Table 8. Results of Gender Diversity as Moderator Between Personality and Acceptance ($k \geq 3$) | NOTE: *Div = Diversity, Grp = Group*
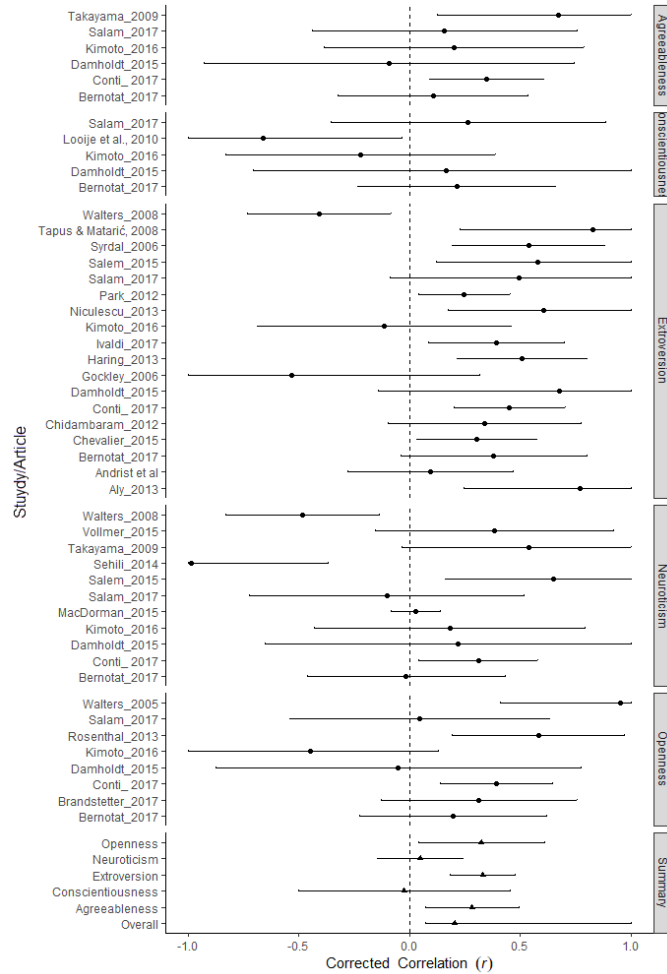
Fig. 3. Forest plot indicates average effect sizes (r) with personality trait as moderator

5.2.3 Gender Diversity. The technology acceptance literature has highlighted the importance of gender differences [97]. Based on these findings we sought to determine whether there is any support for gender differences between sample diversity in terms of personality and acceptance of robots. Results indicated personality was only a predictor of robot acceptance in study samples with high (k=10, $\bar{r}^2$=0.25, 95% CI [0.25,0.48]) and medium (k=4, $\bar{r}^2$=0.34, 95% CI [0.09,0.74]) levels of gender diversity. Likewise personality did not seem to be a predictor in samples with low gender diversity (k=4, $\bar{r}^2$=-0.06, 95% CI [-1.00,0.83]). Specifically, these results indicated a positive correlation between high and medium gender diversity and robot acceptance supporting the idea that samples with more diversity in terms of gender correlate with higher robot acceptance. See Table 8 and Figure 5.

5.2.4 Global Region. To examine the impact of the global region as a moderator, we grouped studies according to where the samples were reportedly collected. Based on the reported information in the articles, data have only been
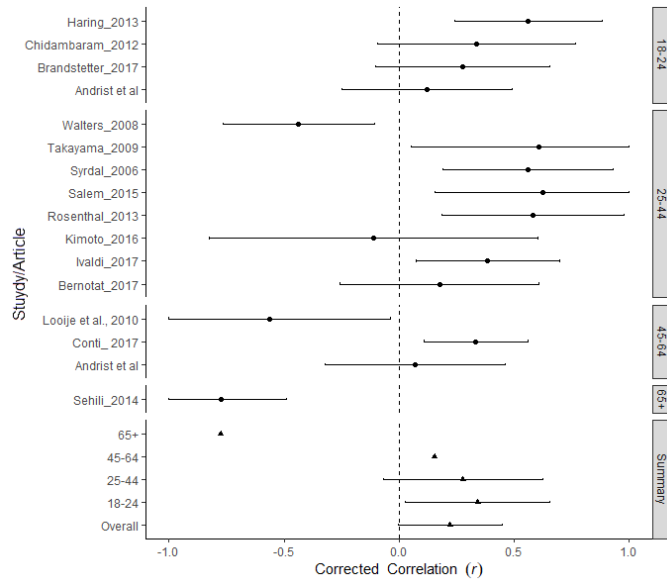
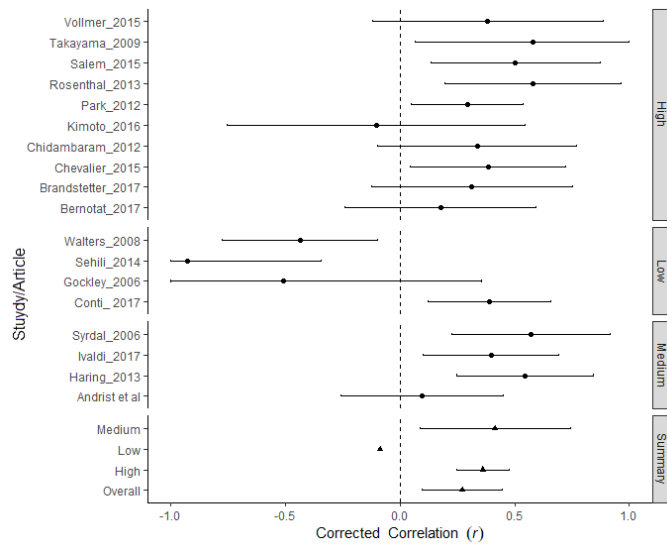Fig. 4. Forrest plot indicates average effect sizes (r) with age group as a moderator

Fig. 5. Forest plot indicates average effect sizes (r) with gender diversity as a moderator

collected from Europe, Asia, Oceania, and North America. As stated, we reported only results with k ≥ 3. Results show that when the effects sizes were divided by the sample region the effects were only significant for the European region. Specifically, results indicate that the positive correlation between personality and robot acceptance was only significant for studies with samples based in Europe. Based on the available data the effects were not significant for Asia or North

15

America. It should be noted that the relationship in the samples based in Asia approached significance unlike the samples based in North America. Details of these results are available in Table 9 and Figure 6. There were not enough studies to examine the impact of the global region by personality trait.

Analysis Type	Predictor	Global Region	Sig	k	N	Mean_r	CI_LL_95	CI_UL_95
Overall	Personality	All Regions	**Y**	18	1437	0.16	0.0559	0.3918
Global Region 1	Personality	Asia	N	3	203	0.31	-0.0653	0.9338
Global Region 2	Personality	Europe	**Y**	10	408	0.32	0.1517	0.7482
Global Region 3	Personality	North America	N	5	826	0.05	-0.2233	0.3561

Table 9. Results of Global Region as Moderator Between Human Personality and Acceptance ($k \geq 3$)
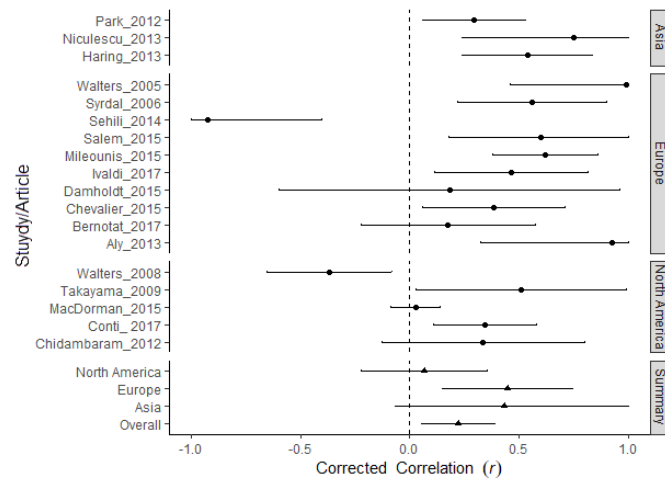

Fig. 6. Forrest plot indicates average effect sizes (r) with global region as moderator

5.2.5   Task Type. Research in HRI has found that different types of tasks are influential to human's perceptions of robots [40, 87]. As a result, we categorized studies according to robot task type based on the available task descriptions in each study. Across studies the coding revealed three broadly defined task types. The three task types were: physical manipulation tasks, influence tasks, and social interaction tasks. Physical manipulation tasks required the robot to physically manipulate the environment or itself. For example, in [53] the robot was asked to retrieve specified items for a participant. Influence required the robot to change the human's attitude or behavior. For example, in [16] the robot was tasked with persuading the participant to select and use certain items over others. Social interaction tasks were those where the robot was required to interact with humans with no specific goal in mind. For example, in [21] the robot was tasked with simply conversing with the participant over lunch.

Task type appeared to moderate the relationship between human personality and robot acceptance. As the results show in Table 10, the data indicated a significant relationship between human personality and robot acceptance for physical manipulation tasks but not for social interaction tasks. In terms of influence tasks, there weren't enough studies to warrant comparison (k=1). Specifically, these results indicate a positive correlation between personality and

Analysis Type	Predictor	Task Type	Sig	k	N	Mean_r	CI_LL_95	CI_UL_95
Overall	Personality	All Task Types	Y	15	702	0.19	0.0418	0.5083
Task Type 1	Personality	Physical Manipulation	Y	5	233	0.3	0.1094	0.7234
Task Type 2	Human Personality	Social/Communication	N	10	469	0.13	-0.1593	0.5415

Table 10. Results of Task Type as Moderator Between Personality and Acceptance ($K \geq 3$)

acceptance for studies using physical manipulation tasks. In terms of different personality traits, comparison was not possible because no single personality trait contained all task types. Details of these results are available in Table 10 and Figure 7.
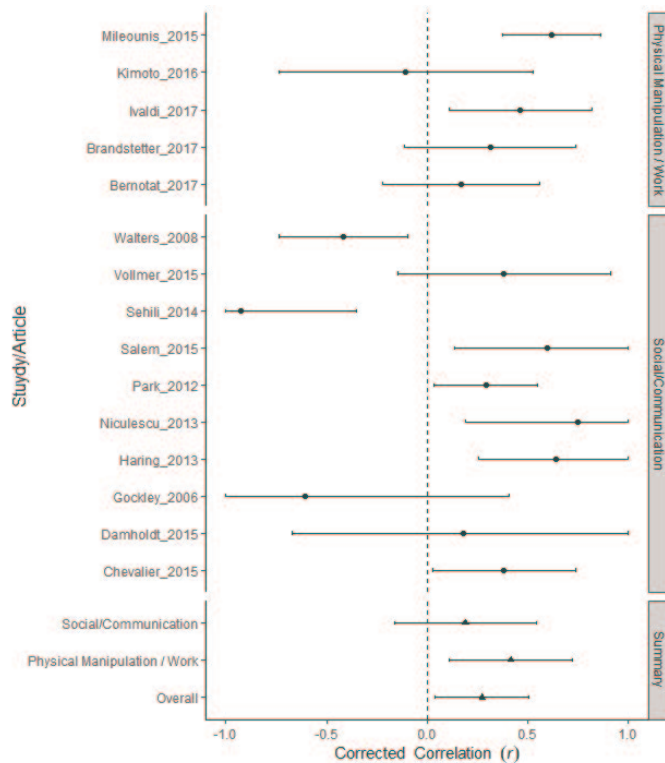

Fig. 7. Forrest plot indicates average effect sizes (r) with task type as a moderator

5.3 Robustness Checks

To assess the overall robustness of our meta-analytical results, we investigated the potential for publication bias in this literature. 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 the results.

5.3.1 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" [85, 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"[7, Pg.123]. Further, "different versions of the fail-safe N produce very different results" [7, Pg.124] which frequently leads to misinterpretation and provides only vague indications of publication bias [7].

The funnel plot analysis (in Figure 8) 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 8 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 present (studies) group 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 is biased toward stronger and positive correlations as opposed to weaker and negative correlations. The degree of bias appears to be relatively moderate because some studies appear in the low to no correlation range. 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.
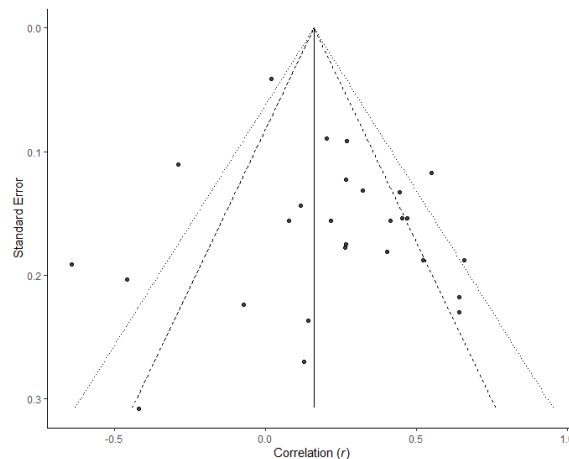

Fig. 8. Funnel plot displays correlation values and standard error

5.3.2 Leave-One-Out Sensitivity Analysis. A leave-one-out sensitivity analysis was performed by iteratively removing one study at a time and running our analysis without that study's data. This analysis provides a sense of robustness and assists in the identification of outliers because we can determine the impact that the exclusion of a study has on our overall findings. Figure 9 and Table 11 present the results of this analysis. Figure 9 presents the mean *ρ* or correlation coefficient of a meta analysis conducted with the named study left out of the analysis. The point is the mean *ρ* bounded by a 95% and 80% confidence interval. The point repeats for each study until the analysis has been run once with all but one of the studies included, for each study. A visual inspection can determine when results shift with the exclusion of a

given study, as was the case with MacDorman (2015) [61]. Table 11 presents these results in a numerical format with the study left out appearing on the far left and the results of a meta analysis excluding it to the right.

As presented in Figure 9 and Table 11, results provide evidence of one potential outlier. This study appears to be [61] because the $/\bar{rho}$ when this study was excluded was higher than when the majority of the other studies were excluded. This indicates that [61] shifted our results toward a negative correlation. This shift was minimal, however, and as a result we opted to keep [61] in our analysis.

Study Left Out	K	N	Mean_r	Sig	CI_LL_95	CI_UL_95
Brandstetter_2017	25	1571	0.142506	Y	0.0556	0.3423
Chevalier_2015	25	1547	0.139338	Y	0.0517	0.3364
Rosenthal_2013	25	1570	0.136325	Y	0.0494	0.3297
Syrdal_2006	25	1570	0.135892	Y	0.0484	0.3288
Gockley_2006	25	1600	0.148254	Y	0.0664	0.344
Haring_2013	25	1556	0.133781	Y	0.0485	0.3261
Walters_2005	25	1583	0.135272	Y	0.0533	0.3254
Niculescu_2013	25	1583	0.137699	Y	0.0518	0.3308
Looije et al., 2010	25	1587	0.153464	Y	0.0759	0.3468
Sehili_2014	25	1584	0.15778	Y	0.0871	0.3458
Vollmer_2015	25	1580	0.142002	Y	0.056	0.3391
Aly_2013	25	1590	0.137786	Y	0.0528	0.3295
Andrist et al	25	1571	0.146088	Y	0.0612	0.3468
Bernotat_2017	25	1564	0.145189	Y	0.0587	0.3437
Chidambaram_2012	25	1579	0.141906	Y	0.0553	0.3397
Conti_ 2017	25	1497	0.134706	Y	0.044	0.3318
Damholdt_2015	25	1597	0.144527	Y	0.0597	0.343
Ivaldi_2017	25	1555	0.137963	Y	0.0512	0.3337
Kimoto_2016	25	1591	0.147101	Y	0.0631	0.3451
MacDorman_2015	25	1040	0.212651	Y	0.1333	0.4567
Park_2012	25	1491	0.139662	Y	0.0482	0.3438
Salam_2017	25	1593	0.1444	Y	0.0593	0.343
Salem_2015	25	1571	0.137477	Y	0.0516	0.3319
Takayama_2009	25	1581	0.13947	Y	0.0522	0.3352
Tapus & Mataric, 2008	25	1592	0.138437	Y	0.0556	0.3312
Walters_2008	25	1532	0.166765	Y	0.0965	0.3612

Table 11. Results of a leave-one-out analysis comparing results with the inclusion or exclusion of a selected study's data.

6 DISCUSSION

The goal of the meta-analysis was to answer three overarching questions. First, does human personality predict robot acceptance? Answer: yes, results of our study identified a strong positive relationship between the various measures of human personality and robot acceptance. Second, are specific personality traits more or less likely to predict robot acceptance? Answer: yes, our results show that extraversion, openness, and agreeableness were significant predictors of robot acceptance while no such evidence was found for conscientiousness or neuroticism (i.e. emotional stability). Third, what role do sample moderators play in this relationship? Moderators had a strong impact on the relationship between human personality and robot acceptance. More specifically, the relationship between human personality varied greatly by the sample's age, gender diversity, task type, and global region. Contributions to the literature, theoretical implications, and study limitations are detailed next.
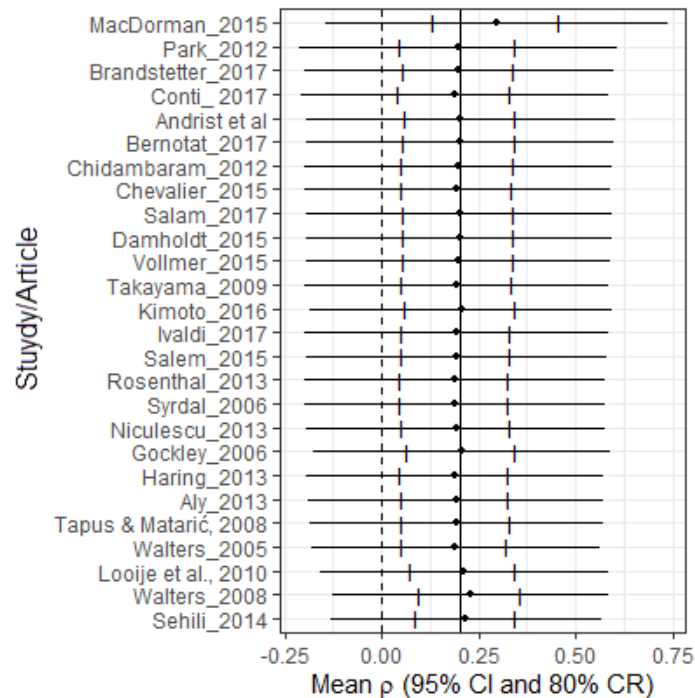
Fig. 9. Leave-one-out analysis results.

6.1 Summary of Findings

First, our results show that the more agreeable, extroverted, and open individuals are the more likely they are to accept a robot. The results of this meta-analysis call attention to the importance of agreeableness, extroversion, and openness on robot acceptance. Every day the importance of robots in our society increases; therefore, the importance of identifying personality traits that are clearly linked to robot acceptance is also increasing. However, no evidence supported the assertion that more or less neurotic or conscientious individuals are any more or less likely to accept a robot. Nonetheless, we are forced to draw very different conclusions from these non-significant results. For neuroticism, the non-significant relationship with acceptance is especially pronounced. Neuroticism had a relatively large sample size and robust study count. Conscientiousness, on the other hand, had a relatively small sample size, making results less conclusive. Although more research is needed to understand the relationship between conscientiousness and acceptance, it is not clear whether such additional studies are warranted for neuroticism.

Second, human personality only significantly predicted robot acceptance in samples of people ages 18–24 years. The relationship between human personality and robot acceptance was non-significant for the other two age groups. One way to interpret this finding is to assume that the relationship between human personality and robot acceptance only truly exists for 18- to 24-year-olds. However, this interpretation may be flawed when we acknowledge that the 45–64

age group only had three studies. Yet, this is not the case for the 25–44 age group, which had twice as many studies as either of the other two age groups. Taken together across all the age groups, it would appear that the relationship between human personality and robot acceptance is most salient for the 18–44 age group. The relationship appears to weaken as the sample age group increases. This finding has several implications for future research. If these results remain consistent, personality should be used cautiously in samples of older age groups. In other words, the older the study sample is, the less likely personality can be used to determine whether participants are likely to accept a robot.

Third, this meta-analysis demonstrated that human personality was only a significant predictor of robot acceptance in samples of high and moderate levels of gender diversity. The relationship between human personality and robot acceptance was not significant for the low-gender-diversity samples. One explanation is that the relationship between personality and robot acceptance differs by actual gender (male vs. female) effects. This would imply that the relationship between personality and robot acceptance would be different in mostly male samples when compared to mostly female samples. To examine this assertion, we conducted an additional analysis by dividing the four low-gender-diversity studies into two groups: two mostly male studies versus two mostly female studies. However, there were no significant differences between the two groups with regard to their relationship between personality and robot acceptance. Therefore, the non-significant results of the low-gender-diversity sample cannot be attributed to gender (i.e., male vs. female) effects.

Fourth, the relationship between human personality and robot acceptance was only significant in studies done in Europe. One way to interpret this finding is that the relationship between human personality and robot acceptance only truly exists for Europeans. However, this interpretation is flawed for several reasons. First, this meta-analysis can only speak to the samples of studies done in Europe. This does not imply that the sample participants themselves consisted entirely of Europeans. In fact, it is unlikely that most studies done with college student samples did not include non-Europeans. Second, this meta-analysis did not include any measure of culture, and assuming that our approach to categorizing the studies into various regions represents culture is at best a crude representation and at worst a gross misrepresentation. Third, the relationship between personality and robot acceptance trended toward significance for studies in Asia. It is quite possible that with additional studies, this relationship would become significant. The non-significant relationship between human personality and robot acceptance in studies done in North America was surprising. This non-significant result was particularly profound because unlike the studies done in Asia, the relationship in the North American samples was near zero despite having more studies than in Asia. Taken together, the findings related to the global region indicate a need for cross-regional studies. Ultimately, the only way to go beyond our initial finding to generalizations regarding nationality or culture is to have additional studies done that specifically examine these issues. The results of this meta-analysis clearly highlight the need for such studies and their potential to contribute to the literature.

6.2 Limitations & Future Research

6.2.1 Additional Studies Needed. The results of this meta-analysis highlight several gaps in the existing literature. First, we see that there is a need for additional studies on both the Big Five personality traits as well as other personality traits. Currently, too few studies have investigated possible moderators within each Big Five personality trait. This made it impossible to examine moderators nested within the Big Five personality traits. If we ever hope to examine possible moderators within each personality trait, more studies are needed. Our study focused exclusively on the Big Five personality traits because there were simply not enough studies on any other personality types. Although the Big Five personality traits are by far the most used traits, other important traits should be examined. For example, the

literature has shown that traits of dominance and honesty-humility are also important predictors of human attitudes and behaviors [76, 78]. Therefore, additional studies are needed to expand the set of personality traits to provide additional insights.

Second, more studies are needed that employ people at a diverse age range. Currently, there is no representation for the 65+ age group in the literature. Practically, when we consider that a large segment of robotic technologies are increasingly targeting health care for those older than 65, we can begin to see an issue with existing sampling (see:[28]). This is an issue because according to table 7, the importance of personality on predicting robot acceptance seems to diminish by age. If this is the case, we may quickly discover that personality is not a useful tool for determining whether individuals older than 65 will accept robots. Therefore, understanding the relationship between personality and robot acceptance for this population is not only under-represented but also timely and important.

Finally, our results highlight a lack of global representation in the research being conducted on human personality and robot acceptance. Unfortunately, personality research in HRI is not being conducted across the globe but instead is focused on a few specific regions: Europe, Asia, and North America. According to our results, there is now reason to solely investigate issues related to the impact of global regions in this research area (see table 9). The lack of representation from other regions is a significant shortcoming. Specifically, it is vital to have studies conducted in Central America and South America as well as the Middle East, North Africa, and sub-Saharan Africa. Until this happens, we will fail to fully understand the relationship between human personality and robot acceptance.

6.2.2 Limitations. This study is not immune to limitations. First, in utilizing the trait-based approach to personality psychology, we are unable to account for inter and intra-subject variability in how people perceive and relate to robots. In addition, there is ongoing debate over the exact number of and definition of each trait examined. In being forced to look at the Big Five personality traits only, we forgo insights that might be visible if other traits were used via scales such as the HEXACO [4] with the addition of *honesty-humility* or the Eysenck Personality Questionnaire [30] which investigates *psychoticism/socialisation.* Furthermore, our study focused on human personality exclusively and does not consider robot personality. We also did not examine issues related to similarity vs. dissimilarity between human and robot personality and the effects this might have because this was outside the scope of our exclusive focus on human personality (see: [81] and [82] for a discussion on the topic). Finally, there is more than one approach to take when considering personality psychology and the HRI literature's emphasis on the trait-based approach overlooks these other schools of thought. Future researchers might wish to adopt other approaches based on, for example, psycho-dynamic or cognitive/social learning perspectives on the subject of human personality.

7 CONCLUSION

This study presents the findings of a meta-analysis on the relationship between human personality and robot acceptance. The results of the study were insightful but produced more questions than answers. This study is an important starting point in our understanding of the relationship between human personality and robot acceptance. Nonetheless, future research is needed to build on these findings and attempt to answer the new questions proposed. Answering these questions will expand our understanding in both an important and much understudied area in the HRI literature.

REFERENCES

[1] Gordon Willard Allport. 1937. Personality: A psychological interpretation. (1937).

[2] Amir Aly and Adriana Tapus. 2013. A model for synthesizing a combined verbal and nonverbal behavior based on personality traits in human-robot interaction. In *2013 8th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. IEEE, 325–332.

[3] Sean Andrist, Bilge Mutlu, and Adriana Tapus. 2015. Look like me: matching robot personality via gaze to increase motivation. In *Proceedings of the 33rd annual ACM conference on human factors in computing systems*. ACM, 3603–3612.

[4] Michael C Ashton, Kibeom Lee, and Reinout E De Vries. 2014. The HEXACO Honesty-Humility, Agreeableness, and Emotionality factors: A review of research and theory. *Pers. Soc. Psychol. Rev.* 18, 2 (2014), 139–152.

[5] Murray R Barrick and Michael K Mount. 1991. The big five personality dimensions and job performance: A meta-analysis. *Pers. Psychol* 44, 1 (1991), 1–26.

[6] Anna Baumert, Manfred Schmitt, Marco Perugini, Wendy Johnson, Gabriela Blum, Peter Borkenau, Giulio Costantini, Jaap JA Denissen, William Fleeson, Ben Grafton, et al. 2017. Integrating personality structure, personality process, and personality development. *European Journal of Personality* 31, 5 (2017), 503–528.

[7] Betsy Jane Becker. 2005. Failsafe N or file-drawer number. *Publication bias in meta-analysis: Prevention, assessment and adjustments* (2005), 111–125.

[8] Jasmin Bernotat and Friederike Eyssel. 2017. A robot at home—How affect, technology commitment, and personality traits influence user experience in an intelligent robotics apartment. In *2017 26th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*. IEEE, 641–646.

[9] LaMarcus R Bolton, Liesl K Becker, and Larissa K Barber. 2010. Big Five trait predictors of differential counterproductive work behavior dimensions. *Personality and Individual Differences* 49, 5 (2010), 537–541.

[10] Michael Borenstein, Larry V Hedges, Julian PT Higgins, and Hannah R Rothstein. 2011. *Introduction to meta-analysis*. John Wiley & Sons.

[11] Gregory J Boyle, Donald H Saklofske, and Gerald Matthews. 2014. *Measures of personality and social psychological constructs*. Academic Press.

[12] Jürgen Brandstetter, Clay Beckner, Eduardo Benitez Sandoval, and Christoph Bartneck. 2017. Persistent lexical entrainment in HRI. In *Proceedings of the 2017 ACM/IEEE International Conference on Human-Robot Interaction*. ACM, 63–72.

[13] Christina Bröhl, Jochen Nelles, Christopher Brandl, Alexander Mertens, and Verena Nitsch. 2019. Human–robot collaboration acceptance model: development and comparison for Germany, Japan, China and the USA. *International Journal of Social Robotics* 11, 5 (2019), 709–726.

[14] David M Buss. 2009. How can evolutionary psychology successfully explain personality and individual differences? *Perspectives on Psychological Science* 4, 4 (2009), 359–366.

[15] Pauline Chevalier, Jean-Claude Martin, Brice Isableu, and Adriana Tapus. 2015. Impact of personality on the recognition of emotion expressed via human, virtual, and robotic embodiments. In *2015 24th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*. IEEE, 229–234.

[16] Vijay Chidambaram, Yueh-Hsuan Chiang, and Bilge Mutlu. 2012. Designing persuasive robots: How robots might persuade people using vocal and nonverbal cues. In *Proceedings of the seventh annual ACM/IEEE international conference on Human-Robot Interaction*. IEEE, 293–300.

[17] Campbell Collaboration et al. 2018. Practical meta-analysis effect size calculator.

[18] Daniela Conti, Elena Commodari, and Serafino Buono. 2017. Personality factors and acceptability of socially assistive robotics in teachers with and without specialized training for children with disability. *Life Span and Disability* 20, 2 (2017), 251–272.

[19] Arturo Cruz-Maya and Adriana Tapus. 2016. Teaching nutrition and healthy eating by using multimedia with a Kompai robot: Effects of stress and user's personality. In *2016 IEEE-RAS 16th International Conference on Humanoid Robots (Humanoids)*. IEEE, 644–649.

[20] Jeffrey A. Dahlke and Brenton M. Wiernik. 2019. psychmeta: An R Package for Psychometric Meta-Analysis. *Applied Psychological Measurement* 43, 5 (2019), 415–416.

[21] Malene F Damholdt, Marco Nørskov, Ryuji Yamazaki, Raul Hakli, Catharina Vesterager Hansen, Christina Vestergaard, and Johanna Seibt. 2015. Attitudinal change in elderly citizens toward social robots: the role of personality traits and beliefs about robot functionality. *Frontiers in psychology* 6 (2015), 1701.

[22] Jonathan J Deeks, Julian PT Higgins, and Douglas G Altman. 2011. 9.5.2 Identifying and measuring heterogeneity. *The Cochrane handbook for systematic reviews of interventions. Version* 5, 0 (2011).

[23] Bryce J Dietrich, Scott Lasley, Jeffery J Mondak, Megan L Remmel, and Joel Turner. 2012. Personality and legislative politics: The Big Five trait dimensions among US state legislators. *Political Psychology* 33, 2 (2012), 195–210.

[24] Andrew Dillon. 2001. User acceptance of information technology. London: Taylor and Francis.

[25] James E Driskell, Gerald F Goodwin, Eduardo Salas, and Patrick Gavan O'Shea. 2006. What makes a good team player? Personality and team effectiveness. *Group Dynamics: Theory, Research, and Practice* 10, 4 (2006), 249.

[26] Yogesh K Dwivedi, Nripendra P Rana, Anand Jeyaraj, Marc Clement, and Michael D Williams. 2019. Re-examining the unified theory of acceptance and use of technology (UTAUT): Towards a revised theoretical model. *Information Systems Frontiers* 21, 3 (2019), 719–734.

[27] Jesse Erdheim, Mo Wang, and Michael J Zickar. 2006. Linking the Big Five personality constructs to organizational commitment. *Personality and individual differences* 41, 5 (2006), 959–970.

[28] Connor Esterwood and Lionel P Robert. 2020. Personality in Healthcare Human Robot Interaction (H-HRI): A Literature Review and Brief Critique. *arXiv preprint arXiv:2008.06723* (2020).

[29] William Fleeson. 2001. Toward a structure-and process-integrated view of personality: Traits as density distributions of states. *Journal of personality and social psychology* 80, 6 (2001), 1011.

[30] Leslie J Francis, Laurence B Brown, and Ronald Philipchalk. 1992. The development of an abbreviated form of the Revised Eysenck Personality Questionnaire (EPQR-A): Its use among students in England, Canada, the USA and Australia. *Personality and individual differences* 13, 4 (1992), 443–449.

[31] Liliane Frey-Rohn. 1974. *From Freud to Jung: A comparative study of the psychology of the unconscious.(Trans FE Engreen & EK Engreen).* GP Putnam's Sons.

[32] David C Funder. 1997. *The personality puzzle.* WW Norton & Co.

[33] Aina Gallego and Sergi Pardos-Prado. 2014. The Big Five personality traits and attitudes towards immigrants. *Journal of Ethnic and Migration Studies* 40, 1 (2014), 79–99.

[34] Alan S Gerber, Gregory A Huber, David Doherty, and Conor M Dowling. 2011. The big five personality traits in the political arena. *Annual Review of Political Science* 14 (2011), 265–287.

[35] Rachel Gockley and Maja J Matarić. 2006. Encouraging physical therapy compliance with a hands-off mobile robot. In *Proceedings of the 1st ACM SIGCHI/SIGART conference on Human-robot interaction.* ACM, 150–155.

[36] Lewis R Goldberg. 1992. The development of markers for the Big-Five factor structure. *Psychological assessment* 4, 1 (1992), 26.

[37] Samuel D Gosling, Peter J Rentfrow, and William B Swann Jr. 2003. A very brief measure of the Big-Five personality domains. *Journal of Research in personality* 37, 6 (2003), 504–528.

[38] Peter A Hancock, Deborah R Billings, Kristin E Schaefer, Jessie YC Chen, Ewart J De Visser, and Raja Parasuraman. 2011. A meta-analysis of factors affecting trust in human-robot interaction. *Human factors* 53, 5 (2011), 517–527.

[39] Kerstin Sophie Haring, Yoshio Matsumoto, and Katsumi Watanabe. 2013. How do people perceive and trust a lifelike robot. In *Proceedings of the world congress on engineering and computer science*, Vol. 1.

[40] Markus Häring, Dieta Kuchenbrandt, and Elisabeth André. 2014. Would You Like to Play with Me? How Robots' Group Membership and Task Features Influence Human–Robot Interaction. In *2014 9th ACM/IEEE International Conference on Human-Robot Interaction (HRI).* IEEE, 9–16.

[41] Anne W Harzing. 2007. Publish or Perish App. RetreivedNov.01,2020fromhttps://harzing.com/resources/publish-or-perish

[42] Nick Haslam. 2007. *Trait Psychology.* 1, Vol. 1. SAGE Publications Ltd, Chapter 2, 17–45.

[43] Larry Hedges and Ingram Olkin. 2014. *Statistical methods for meta-analysis.* Academic press.

[44] Marcel Heerink, Ben Krose, Vanessa Evers, and Bob Wielinga. 2009. 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, 528–533.

[45] JPT Higgins and S Green. 2008. How to include multiple groups from one study. *Cochrane handbook for systematic reviews of interventions version* 5, 0 (2008).

[46] Julian PT Higgins, Simon G Thompson, Jonathan J Deeks, and Douglas G Altman. 2003. Measuring inconsistency in meta-analyses. *Bmj* 327, 7414 (2003), 557–560.

[47] Kasper Hornbæk and Effie Lai-Chong Law. 2007. Meta-analysis of correlations among usability measures. In *Proceedings of the SIGCHI conference on Human factors in computing systems.* ACM, 617–626.

[48] John E Hunter and Frank L Schmidt. 2004. *Methods of meta-analysis: Correcting error and bias in research findings.* Sage.

[49] Serena Ivaldi, Sebastien Lefort, Jan Peters, Mohamed Chetouani, Joelle Provasi, and Elisabetta Zibetti. 2017. Towards engagement models that consider individual factors in HRI: On the relation of extroversion and negative attitude towards robots to gaze and speech during a human–robot assembly task. *International Journal of Social Robotics* 9, 1 (2017), 63–86.

[50] Oliver P John, Laura P Naumann, and Christopher J Soto. 2008. Paradigm shift to the integrative big five trait taxonomy. *Handbook of personality: Theory and research* 3, 2 (2008), 114–158.

[51] Oliver P John, Sanjay Srivastava, et al. 1999. The Big Five trait taxonomy: History, measurement, and theoretical perspectives. *Handbook of personality: Theory and research* 2, 1999 (1999), 102–138.

[52] Hager Khechine, Sawsen Lakhal, and Paterne Ndjambou. 2016. A meta-analysis of the UTAUT model: Eleven years later. *CJAS* 33, 2 (2016), 138–152.

[53] Mitsuhiko Kimoto, Takamasa Iio, Masahiro Shiomi, Ivan Tanev, Katsunori Shimohara, and Norihiro Hagita. 2016. Relationship between personality and robots' interaction strategies in object reference conversations. In *Proceedings of the Second International Conference on Electronics and Software Science (ICESS2016), Japan.* 128–136.

[54] Daniël Lakens. 2013. Calculating and reporting effect sizes to facilitate cumulative science: a practical primer for t-tests and ANOVAs. *Frontiers in psychology* 4 (2013), 863.

[55] Kwan Min Lee and Clifford Nass. 2003. Designing social presence of social actors in human computer interaction. In *Proceedings of the SIGCHI conference on Human factors in computing systems.* ACM, 289–296.

[56] Wolgang Lenhard and Alexandra Lenhard. 2016. Calculation of effect sizes.

[57] Ning Li, Murray R Barrick, Ryan D Zimmerman, and Dan S Chiaburu. 2014. Retaining the productive employee: The role of personality. *Academy of Management Annals* 8, 1 (2014), 347–395.

[58] Mark W Lipsey and David B Wilson. 2001. *Practical meta-analysis.* Sage publications, Inc.

[59] Rosemarijn Looije, Mark A Neerincx, and Fokie Cnossen. 2010. Persuasive robotic assistant for health self-management of older adults: Design and evaluation of social behaviors. *International Journal of Human-Computer Studies* 68, 6 (2010), 386–397.

[60] Tengteng Ma, Hasti Sharifi, and Debaleena Chattopadhyay. 2019. Virtual Humans in Health-Related Interventions: A Meta-Analysis. In *Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems.* 1–6.

[61] Karl F MacDorman and Steven O Entezari. 2015. Individual differences predict sensitivity to the uncanny valley. *Interaction Studies* 16, 2 (2015), 141–172.

[62] Gerald Matthews, Ian J Deary, and Martha C Whiteman. 2003. *Personality traits*. Cambridge University Press.

[63] Robert R McCrae and Paul T Costa. 2003. *Personality in adulthood: A five-factor theory perspective*. Guilford Press.

[64] Robert R McCrae and Paul T Costa Jr. 1997. Personality trait structure as a human universal. *American psychologist* 52, 5 (1997), 509.

[65] Robert R McCrae and Paul T Costa Jr. 2008. The five-factor theory of personality. (2008).

[66] Jim McMartin. 2016. *Personality psychology: A student-centered approach*. Sage Publications.

[67] Antonia Meissner, Angelika Trübswetter, Antonia S Conti-Kufner, and Jonas Schmidtler. 2020. Friend or foe? understanding assembly workers' acceptance of human-robot collaboration. *ACM Transactions on Human-Robot Interaction (THRI)* 10, 1 (2020), 1–30.

[68] Taciano L Milfont and Chris G Sibley. 2012. The big five personality traits and environmental engagement: Associations at the individual and societal level. *Journal of Environmental Psychology* 32, 2 (2012), 187–195.

[69] Walter Mischel. 1973. Toward a cognitive social learning reconceptualization of personality. *Psychological review* 80, 4 (1973), 252.

[70] Walter Mischel. 1990. Personality dispositions revisited and revised: A view after three decades. (1990).

[71] Tatsuya Nomura, Takayuki Kanda, Tomohiro Suzuki, and Kensuke Kato. 2008. Prediction of human behavior in human–robot interaction using psychological scales for anxiety and negative attitudes toward robots. *IEEE transactions on robotics* 24, 2 (2008), 442–451.

[72] Tatsuya Nomura, Takuya Shintani, Kazuki Fujii, and Kazumasa Hokabe. 2007. Experimental investigation of relationships between anxiety, negative attitudes, and allowable distance of robots. In *Proceedings of the 2nd IASTED international conference on human computer interaction, Chamonix, France. ACTA Press*. 13–18.

[73] Kohei Ogawa, Christoph Bartneck, Daisuke Sakamoto, Takayuki Kanda, Tetsuo Ono, and Hiroshi Ishiguro. 2018. Can an android persuade you? In *Geminoid Studies*. Springer, 235–247.

[74] Eunil Park, Dallae Jin, and Angel P del Pobil. 2012. The law of attraction in human-robot interaction. *International Journal of Advanced Robotic Systems* 9, 2 (2012), 35.

[75] Miranda AG Peeters, Harrie FJM Van Tuijl, Christel G Rutte, and Isabelle MMJ Reymen. 2006. Personality and team performance: a meta-analysis. *European Journal of Personality: Published for the European Association of Personality Psychology* 20, 5 (2006), 377–396.

[76] Sofia Petisca, Francisco Esteves, and Ana Paiva. 2019. Cheating with robots: how at ease do they make us feel?. In *IROS*. 2102–2107.

[77] Kym E Pocius. 1991. Personality factors in human-computer interaction: A review of the literature. *Computers in Human Behavior* 7, 3 (1991), 103–135.

[78] Aaron Powers and Sara Kiesler. 2006. 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*. ACM, 218–225.

[79] Natalia Reich-Stiebert and Friederike Eyssel. 2015. Learning with educational companion robots? Toward attitudes on education robots, predictors of attitudes, and application potentials for education robots. *International Journal of Social Robotics* 7, 5 (2015), 875–888.

[80] Jinny Rhee, David Parent, and Anuradha Basu. 2013. The influence of personality and ability on undergraduate teamwork and team performance. *SpringerPlus* 2, 1 (2013), 16.

[81] Lionel Robert. 2018. Personality in the human robot interaction literature: A review and brief critique. In *Proceedings of the 24th Americas Conference on Information Systems, Aug.* 16–18.

[82] Lionel P. Robert., Rasha Alahmad, Connor Esterwood, Sangmi Kim, Sangseok You, and Qiaoning Zhang. 2020. A Review of Personality in Human–Robot Interactions. *Foundations and Trends® in Information Systems* 4, 2 (2020), 107–212.

[83] Lionel P Robert Jr and Tracy Ann Sykes. 2017. Extending the concept of control beliefs: Integrating the role of advice networks. *Information Systems Research* 28, 1 (2017), 84–96.

[84] Astrid M Rosenthal-von der Pütten, Nicole C Krämer, Laura Hoffmann, Sabrina Sobieraj, and Sabrina C Eimler. 2013. An experimental study on emotional reactions towards a robot. *International Journal of Social Robotics* 5, 1 (2013), 17–34.

[85] Hannah R Rothstein, Alexander J Sutton, and Michael Borenstein. 2005. Publication bias in meta-analysis. *Publication bias in meta-analysis: Prevention, assessment and adjustments* (2005), 1–7.

[86] Hanan Salam, Oya Celiktutan, Isabelle Hupont, Hatice Gunes, and Mohamed Chetouani. 2016. Fully automatic analysis of engagement and its relationship to personality in human-robot interactions. *IEEE Access* 5 (2016), 705–721.

[87] Maha Salem, Gabriella Lakatos, Farshid Amirabdollahian, and Kerstin Dautenhahn. 2015. Would you trust a (faulty) robot? Effects of error, task type and personality on human-robot cooperation and trust. In *2015 10th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. IEEE, 1–8.

[88] Neil J Salkind. 2010. *Encyclopedia of research design*. Vol. 1. Sage. 423–424 pages.

[89] Gerard Saucier. 2008. Measures of the personality factors found recurrently in human lexicons. *Handbook of personality theory and testing* 2 (2008), 29–54.

[90] Mohamed Sehili, Fan Yang, Violaine Leynaert, and Laurence Devillers. 2014. A corpus of social interaction between nao and elderly people. In *5th International Workshop on Emotion, Social Signals, Sentiment & Linked Open Data (ES3LOD2014). LREC*.

[91] Chris G Sibley, Danny Osborne, and John Duckitt. 2012. Personality and political orientation: Meta-analysis and test of a Threat-Constraint Model. *Journal of Research in Personality* 46, 6 (2012), 664–677.

[92] Dag Sverre Syrdal, Kerstin Dautenhahn, Sarah Woods, Michael L Walters, and Kheng Lee Koay. 2006. 'Doing the right thing wrong'-Personality and tolerance to uncomfortable robot approaches. In *ROMAN 2006-The 15th IEEE International Symposium on Robot and Human Interactive Communication*. IEEE, 183–188.

[93] Leila Takayama and Caroline Pantofaru. 2009. Influences on proxemic behaviors in human-robot interaction. In *2009 IEEE/RSJ International Conference on Intelligent Robots and Systems*. IEEE, 5495–5502.

[94] Kevin Tasa, Greg J Sears, and Aaron CH Schat. 2011. Personality and teamwork behavior in context: The cross-level moderating role of collective efficacy. *Journal of Organizational Behavior* 32, 1 (2011), 65–85.

[95] Will Thalheimer and Samantha Cook. 2002. How to calculate effect sizes from published research: A simplified methodology. *Work-Learning Research* 1 (2002), 1–9.

[96] Carl J Thoresen, Seth A Kaplan, Adam P Barsky, Christopher R Warren, and Kelly De Chermont. 2003. The affective underpinnings of job perceptions and attitudes: a meta-analytic review and integration. (2003).

[97] Viswanath Venkatesh and Michael G Morris. 2000. Why don't men ever stop to ask for directions? Gender, social influence, and their role in technology acceptance and usage behavior. *MISQ* (2000), 115–139.

[98] Viswanath Venkatesh, Michael G Morris, Gordon B Davis, and Fred D Davis. 2003. User acceptance of information technology: Toward a unified view. *MISQ* (2003), 425–478.

[99] Sarah Theres Völkel, Ramona Schödel, Daniel Buschek, Clemens Stachl, Quay Au, Bernd Bischl, Markus Bühner, and Heinrich Hussmann. 2019. 2 Opportunities and challenges of utilizing personality traits for personalization in HCI. *Personalized Human-Computer Interaction* (2019), 31.

[100] Anna-Lisa Vollmer, Katharina J Rohlfing, Britta Wrede, and Angelo Cangelosi. 2015. Alignment to the Actions of a Robot. *International Journal of Social Robotics* 7, 2 (2015), 241–252.

[101] Michael L Walters, Kerstin Dautenhahn, René Te Boekhorst, Kheng Lee Koay, Christina Kaouri, Sarah Woods, Chrystopher Nehaniv, David Lee, and Iain Werry. 2005. The influence of subjects' personality traits on personal spatial zones in a human-robot interaction experiment. In *ROMAN 2005. IEEE International Workshop on Robot and Human Interactive Communication, 2005*. IEEE, 347–352.

[102] Michael L Walters, Dag S Syrdal, Kerstin Dautenhahn, René Te Boekhorst, and Kheng Lee Koay. 2008. Avoiding the uncanny valley: robot appearance, personality and consistency of behavior in an attention-seeking home scenario for a robot companion. *Autonomous Robots* 24, 2 (2008), 159–178.

[103] Astrid Weiss, Betsy Van Dijk, and Vanessa Evers. 2012. Knowing me knowing you: Exploring effects of culture and context on perception of robot personality. In *Proceedings of the 4th international conference on Intercultural Collaboration*. ACM, 133–136.

[104] Martin J Westgate. 2019. revtools: An R package to support article screening for evidence synthesis. *Research synthesis methods* (2019).

[105] Nick Yee, Jeremy N Bailenson, and Kathryn Rickertsen. 2007. A meta-analysis of the impact of the inclusion and realism of human-like faces on user experiences in interfaces. In *Proceedings of the SIGCHI conference on Human factors in computing systems*. ACM, 1–10.

[106] Sangseok You and Lionel Robert. 2018. Teaming up with robots: An IMOI (inputs-mediators-outputs-inputs) framework of human-robot teamwork. *Int.J Robotic Eng(IJRE)* 2, 3 (2018).