Original Papers

Polish Psychological Bulletin
2025, vol. 56 148–160
DOI: 10.24425/ppb.2025.153987

Katarzyna Samson, Tomasz Zaleskiewicz

SWPS University, Wroclaw, Poland

Corresponding author – Katarzyna Samson – kasiasamson@gmail.com

Warmth and Competence in Human-AI Agent Interactions

Abstract: As artificial intelligence (AI) technologies increasingly integrate into daily life, understanding how people perceive AI agents in trust-related interactions is critical for fostering effective human-AI collaboration. Drawing on social cognition theory, this research examines the fundamental dimensions of warmth and competence in shaping impressions and trust towards AI agents compared to humans. Across two studies using trust-related vignettes, we investigated how warmth and competence are attributed to AI agents, human experts, and friends in various social contexts and performance outcomes. The results indicated that although both warmth and competence impact trust judgments, these traits are generally considered less important and are attributed to AI agents to a lesser extent than to humans. Moreover, AI agents were perceived as equally warm and competent, whereas humans were rated higher on both traits—especially on warmth. The findings highlight the nuanced role of social cognitive dimensions in human-AI trust, suggesting that perceptions of AI are context-dependent and affected by implicit biases. This work advances understanding of human-AI social dynamics and underscores the importance of designing AI systems that effectively balance warmth and competence to enhance trust and cooperation.

Keywords: *warmth, competence, social cognition, trust, artificial intelligence*

With the development of artificial intelligence (AI), a human-AI society is becoming a reality. AI technologies are increasingly common and continue to evolve, integrating more deeply into all areas of life in various forms and serving a wide range of functions. AI agents now operate as autonomous, independent entities across diverse contexts, offering benefits for economic growth, productivity, and safety in multiple industries (Espina-Romero et al., 2023; Malik et al., 2024). These technologies are no longer seen merely as tools; people perceive and interact with them as if they were real social actors (Harris-Watson et al., 2023; McKee et al., 2023; Nass et al., 1994; Reeves & Nass, 1996). To facilitate harmonious relationships in a human-AI society, it is essential to understand how people perceive and form their impressions of AI agents.

Research on social cognition (Fiske, 1993, 2018; Fiske et al. 2002, 2007) offers a valuable framework for understanding human-AI relationships. The two fundamental dimensions of social cognition are warmth, which reflects whether a person's intentions are benevolent or malevolent, and competence, reflecting their ability to act on those intentions. Perceptions of an interaction partner's warmth and competence are linked to specific emotional and behavioral responses. Individuals perceived as both warm and competent evoke pride and admiration; those seen as warm but incompetent elicit pity and sympathy; those viewed as cold but competent trigger envy and jealousy; and, finally, those perceived as cold and incompetent provoke disgust and contempt (Fiske et al., 2002; Fiske, 2018).

While both warmth and competence are fundamental to social perception, evidence suggests that warmth judgments typically precede competence judgments and carry greater weight in shaping emotional and behavioral reactions (Fiske, 2018; Fiske et al., 2007). This primacy of warmth appears to have an evolutionary basis: recognizing whether others intend to help or harm us may be more critical for our survival than assessing their ability to do so.

This work was funded by Narodowe Centrum Nauki grant no: 2019/35/D/HS6/03113 awarded to KS.

In consequence, interpersonal trust tends to depend more on perceptions of warmth than on perceptions of competence (Fiske, 2018; Fiske et al., 2007).

These social cognitive judgments extend beyond human actors. Warmth and competence have also been identified as meaningful dimensions in perceptions of AI agents (Bergmann et al., 2012; Harris-Watson et al. 2023; Khadpe et al., 2020; McKee et al., 2023). For AI, warmth corresponds to perceived human-AI interdependence (aligned interests), while competence is linked to perceived autonomy and effectiveness (McKee et al., 2023). AI systems that align with human interests are seen as warmer, whereas those operating independently are perceived as more competent (McKee et al., 2023). Nonetheless, technical systems and AI agents are generally perceived as both less warm and less competent than humans (Ashktorab et al., 2020; Frischknecht, 2021; van der Hoeven, 2018). Additionally, AI systems are often regarded as more competent than warm (McKee et al., 2023). However, the precise role of warmth and competence in trust formation within human-AI interactions remains unclear. While both dimensions influence trust in AI systems (Li et al., 2025) and robots (Christoforakos et al., 2021), findings are mixed: some meta-analyses highlight competence as more impactful on trust (Hancock et al., 2011), while other studies emphasize the role of warmth (Gilad et al. 2021; Kulms & Kopp, 2018). These discrepancies suggest that the relative importance of warmth and competence may differ between human-AI and human-human trust interactions.

PRESENT RESEARCH

In the present research, we directly compared trust-related interactions between humans and AI agents through the lens of two fundamental dimensions of social cognition: warmth and competence. Given the wide array of situations described across the experimental materials, we chose not to narrow the general category of automated systems. Instead, we refer to all non-human social actors as "AI agents," allowing participants to mentally substitute this broad term with whatever technology they find most familiar and appropriate for each situation.

In Study 1, participants evaluated the importance of these traits in everyday trust scenarios involving either human or AI trustees, using vignette-based descriptions. We hypothesized that the importance of warmth and competence would vary depending on both the type of trustee and the situation. Warmth would be more important for human trustees and in scenarios where people would prefer to consult a friend rather than an AI agent, whereas competence would be more important for AI agent trustees and in scenarios where people would be more inclined to consult an AI agent.

Study 2 expanded on this by examining how warmth and competence were attributed when explaining the success or failure of trust interactions involving one of three trustee types: an AI agent, a human expert, or a human friend. We hypothesized that warmth and competence attributions would vary across the three trustee types as a function of performance: overall, human trustees would be attributed more warmth, while AI agent trustees would be attributed more competence. We further hypothesized that AI agents would be judged more harshly, such that failure would reduce their perceived warmth and competence more than it would for human trustees.

GENERAL METHODS

Materials

In both studies, we used the same set of six vignettes selected from a pool of 15 vignettes that we pretested in a pilot study (N = 200, age M = 32.25, SD = 11.24, 55% Female, 45% Male). The pilot study assessed participants' preferences for interacting with (a) an expert or a friend; and (b) a human or an AI agent. The results identified three types of situations: (1) those in which people preferred to trust a friend, (2) those in which people preferred to trust a human expert, and (3) those in which people preferred to trust an AI agent.

To ensure our research captured a broad spectrum of trust-related scenarios without bias towards any particular type, we selected two vignettes from each scenario type (friend preference, expert preference, AI agent preference) where preferences were most pronounced. The pilot study results are presented in the Appendix. The six vignettes used in the main research are listed below.

a) You take care of your family member and you need someone to take over your duties while you're away for a couple of days [a friend preference type].
b) You had a fight with a close friend and you need advice regarding how to handle it [a friend preference type].
c) You have a legal problem that could lead to serious consequences and you don't know how to handle it [a human expert preference type].
d) You have a non-life-threatening health issue and you don't know what to do about it [a human expert preference type].
e) You received a significant amount of extra money and you don't know how to invest it [an AI agent preference type].
f) You've come across a technical, work-related issue and you're looking for a way to address it [an AI agent preference type].

Sample Sizes

Sample sizes in presented studies were determined based on a rule of thumb of 100 participants per experimental condition. Cohen (1988) recommends a sample size of approximately 64 participants per condition to achieve a power of .80 with a significance level of 0.05 for a medium-sized effect of d = 0.50, and larger samples for smaller effect sizes. However, in social psychology effect sizes tend to be smaller, the average effect being approximately d = .40, (Camerer et al., 2018; Lovakov & Agadullina, 2001; Richard et al., 2003). Therefore, given the smaller expected effect size, we increased the sample sizes relative to Cohen's recommendations.

Normality Assumption

The normality assumption for the dependent variables was not met in either of the studies, as indicated by the Kolmogorov-Smirnov test (results are presented in the Appendix). However, the fixed effects model F-test is robust to deviations from normality, especially if the sample sizes are equal and large (Blanca, 2017; Osborne, 2008). Moreover, the skewness of the distribution – more important in terms of inflating the probability of type I error (Osborne, 2008) – was within the rule of thumb range of (-2, +2) for each variable, so we decided the F-test would be a valid choice despite the violated normality assumption.

STUDY 1

In Study 1, we examined the importance of warmth and competence of the trustee in situations requiring trust. Using the experimental vignette method, we measured the perceived importance of trustee characteristics (warmth and competence) depending on trustee type (AI agent or human) and situation type (friend preference, expert preference, or AI agent preference).

We hypothesized that the importance of warmth and competence of the trustee would vary by trustee type, such that (1) in the case of human trustees, warmth would be more important in trust-related decisions, while (2) in the case of AI agent trustees, competence would be more important. We also hypothesized that the importance of warmth and competence of the trustee would vary by situation type, such that (3) in the friend-preference scenarios warmth would be more important than in the expert-preference and the AI agent-preference scenarios, while (4) in the expert-preference and the AI agent-preference scenarios, competence would be more important than in the friend-preference scenarios.

Participants

The study was conducted on an online sample (N = 198; 45.2% female, 54.8% male; age 18-80 M = 30.68, SD = 9.31) recruited via the Prolific platform. Participants were compensated £0.45, corresponding to an hourly rate of £9.

There was no significant difference in gender distribution across conditions, $\chi^2(1, N = 198) = 0.88$, $p = .35$. There were also no significant differences in age distribution across conditions, $t(196) = -0.96$, $p = .34$.

Procedure

Participants were asked to imagine themselves in each of the six situations presented in the vignettes (two of each of the three situation types) in which they considered seeking advice or support from the trustee. They did not interact directly with the trustee. After reading each vignette, participants rated the importance of trustee characteristics (first warmth, then competence) for deciding whether to trust them. Responses were given on a Likert-type scale ranging from 1–*not at all* to 7–*extremely*. The order of presented situations was randomized.

Measures

To measure the importance of warmth, we used the following item: "It's important that the [person/AI agent] is friendly".

To measure the importance of competence, we used the following item: "It's important that the [person/AI agent] is competent".

Design

The study employed a 2(trustee type: between-participants) x 2(trustee characteristics: within-participants) x 3(situation type: within-participants) experimental design.

Results

Results – Warmth and Competence

Overall, the correlation between warmth and competence scores was large, $r(197) = 0.55$, $p < .001$. We also computed correlations between warmth and competence scores for each situation type. For the AI agent-related scenarios, the correlation between warmth and competence was $r(197) = 0.28$, $p < .001$, CI [0.15, 0.40]. For the expert-related scenarios, the correlation between warmth and competence was $r(197) = 0.34$, $p < .001$, CI [0.21, 0.46]. For the friend-related scenarios, the correlation between warmth and competence was $r(197) = 0.68$, $p < .001$, CI [0.60, 0.75].

Results – Hypotheses Testing

We conducted a three-way mixed ANOVA with trustee type (AI agent vs. human) as a between-participants factor and situation type (AI agent vs. expert vs. friend) and trustee characteristics (warmth vs. competence) as within-participants factors (see Fig. 1).

Mauchly's test indicated that the assumption of sphericity was violated for situation type, $\chi^2(2, N = 198) = 72.05$, $p < .001$, and for situation type by trustee characteristics interaction, $\chi^2(2, N = 198) = 37.87$, $p < .001$. Therefore, degrees of freedom were corrected using the Huynh-Feldt estimate of sphericity, $\Sigma = .76$ and $\Sigma = .85$, respectively.

The results showed a significant main effect of trustee type, indicating that the importance of trustee characteristics was generally higher for human trustees (M = 5.51, SE = 0.12) than for AI agent trustees (M = 5.14, SE = 0.12), $F(1,196) = 4.82$, $p = .03$, $\eta^2_p = .02$. The main effect of trustee characteristics was also significant, $F(1,196) = 117.12$, $p < .001$, $\eta^2_p = .37$, indicating that competence (M = 5.82, SE = 0.09) was generally more important than warmth (M = 4.83, SE = 0.10). The main effect of the situation type was not significant, $F(1.55, 302.90) = 1.02$, $p = .35$, $\eta^2_p = .01$.

The interaction between situation type and trustee characteristics was significant, $F(1.72, 337.51) = 124.12$, $p < .001$, $\eta^2_p = .39$. The interactions between trustee type and situation type, as well as trustee type and trustee characteristics, were not significant, $F(1.55, 302.90) = 1.64$, $p = .20$, $\eta^2_p = .01$ and $F(1, 196) = 1.25$, $p = .27$,

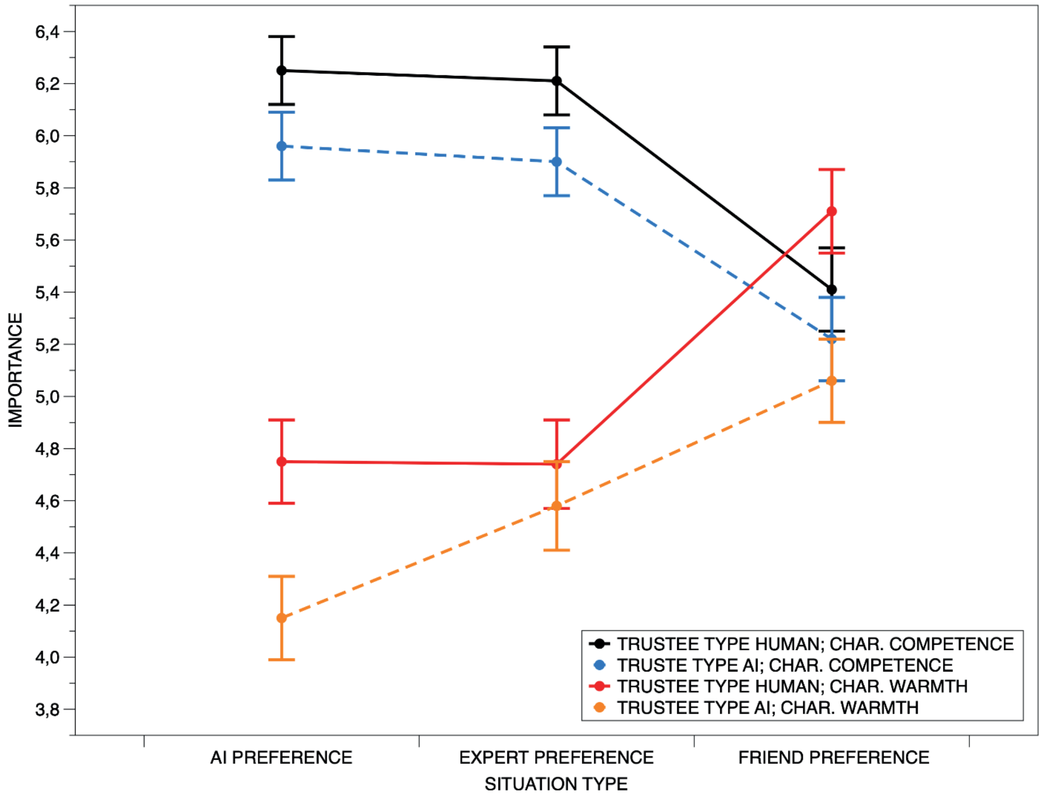

Figure 1. *The Importance of the Trustee Characteristics– Warmth (Red and Orange Lines) and Competence (Black and Blue Lines) – for Human (Solid Line) and AI Trustees (Dashed Line) as a Function of Situation Type.*

Note. Error bars represent 95% confidence intervals.

η^2_p = .01, respectively. The three-way interaction among trustee type, trustee characteristics, and situation type was significant, $F(1.72, 337.51) = 3.70$, $p = .03$ $\eta^2_p = .02$.

Our hypotheses – that (1) warmth would be more important than competence when evaluating humans, and (2) competence would be more important than warmth when evaluating AI agent trustees – were not supported, as the interaction between trustee type and trustee characteristics was not significant.

To test Hypotheses (3) and (4) – that the importance of warmth and competence would vary depending on the situation type – we conducted pairwise comparisons for the situation type x trustee characteristics interaction, using the Bonferroni adjustment. The importance of warmth significantly differed by the type of a situation, $F(2,195) = 35.34$, $p < .001$, $\eta^2_p = .27$. In friend-related scenarios (M = 5.39, SE = 0.11), warmth was rated as more important than in expert-related scenarios (M = 4.66, SE = 0.12), p < .001, and AI agent-related scenarios (M = 4.45, SE = 0.11), p < .001, supporting our Hypothesis (3). The importance of competence also differed by situation type, $F(2,195) = 46.40$, $p < .001$, $\eta^2_p = .32$. In friend-related scenarios (M = 5.31, SE = 0.12), competence was rated as less important than in expert-related scenarios (M = 6.05, SE = 0.09), $p < .001$, and AI agent-related scenarios (M = 6.10, SE = 0.09), $p < .001$– supporting our Hypothesis (4). There was no significant difference in the importance of competence between expert-related scenarios and AI agent-related scenarios, $p = .90$, but the warmth dimension was rated slightly higher in expert-related scenarios than in AI agent-related scenarios, $p = .02$.

In summary, results from a mixed ANOVA revealed that trustee characteristics were considered more important when evaluating human trustees than AI agents. Overall, competence was rated as more important than warmth. These ratings varied by situation type: warmth was seen as more important in friend-related scenarios, while competence was prioritized in situations involving experts and AI agents. Contrary to our hypotheses, there was no significant interaction between trustee type and trustee characteristics—warmth was not uniquely emphasized for humans, nor competence for AI agents. However, a three-way interaction was also found between trustee type, trustee characteristic, and situation type. We also found that warmth and competence scores are more correlated in case of friend-related scenarios than in case of AI agent- or expert-related scenarios.

STUDY 2

In the second study, we examined how people retrospectively explain trust-related situations – i.e., when they already know their outcome. Using the experimental vignette method, we measured attributions of two trustee characteristics (warmth and competence) as a function of trustee performance (success or failure) and trustee type (AI agent or expert or friend).

We hypothesized that warmth and competence attributions would vary across the three trustee types depending on trustee performance in such way that (5) friends would have higher warmth scores than experts and experts would have higher warmth scores than AI agents, while (6) AI agents would have higher competence scores than experts, and experts would have higher competence scores than friends. We also hypothesized that (7) the effect of trustee performance would be stronger in the case of AI agent trustees than in the case of experts and friends.

Participants

The second study was conducted on an online sample (N = 600; 40.2% male, 59.3% female, 0.5% other; age 18-80 M = 31.44, SD = 10.13) recruited via the Prolific platform. Participants were compensated £0.45, corresponding to an hourly rate of £9.

There was no significant difference in gender distribution across conditions, χ^2 (10, N = 598) = 4.25, p = .94. There were also no significant differences in age across conditions, F(5,598) = .67, p = .64.

Procedure

Participants were first asked to imagine themselves in each of the six presented situations. They did not interact directly with the trustee but were provided a brief description of each hypothetical scenario. After reading each scenario, participants were informed of the trustee's performance (success or failure) and asked to rate the trustee on warmth and competence (in fixed order), using a Likert-type scale ranging from 1–*not at all* to 7–*extremely*. The order of presented scenarios was randomized.

Below is a sample vignette illustrating the trustee type and performance manipulations: „You have a legal problem that could lead to serious consequences and you don't know how to handle it. You decide to consult a/an [AI agent/expert/friend]. You follow their advice and end up really [happy/unhappy] about the consequences".

Measures

To measure the trustee's warmth, we used the following item: "In your opinion, the [AI agent/expert/friend] is friendly."

To measure the trustee's competence, we used the following item: "In your opinion, the [AI agent/expert/friend] is competent."

Design

The study employed a 3(trustee type: between-participants) x 2(trustee performance: between-participants) x 2(trustee characteristics: within-participants).

Results

Results – Warmth and Competence

Overall, the correlation between warmth and competence scores was large, r(599) = 0.76, p < .001. We also computed correlations between warmth and competence scores for each situation type. For the AI agent-related scenarios, the correlation between warmth and competence was r(599) = 0.68, p < .001, CI [0.64, 0.73]. For the expert-related scenarios, the correlation between warmth and competence was r(599) = 0.69, p < .001, CI [0.65, 0.73]. For the friend-related scenarios, the correlation between warmth and competence was r(599) = 0.79, p < .001, CI [0.76, 0.82].

Results – Scenario Type

Warmth scores were strongly intercorrelated across the six scenarios (Cronbach's α = .96), and competence scores were similarly intercorrelated (Cronbach's α = .98), so we averaged warmth and competence scores across all scenarios.

Results – Hypotheses testing

We conducted a three-way mixed ANOVA with trustee type (AI agent vs. expert vs. friend) and trustee performance (success vs. failure) as between-participants factors and trustee characteristics (warmth vs competence) as a within-participants factor (see Fig. 2).

With only one two-level within-subject factor the assumption of sphericity was automatically met, so no correction for degrees of freedom was applied. The results showed significant main effects of trustee type and trustee performance, F(2,594) = 51.05, p < .001, η^2_p = .15 and F(1,594) = 1147.73, p < .001, η^2_p = .66, respectively. Trustee characteristics ratings differed across the three trustee types (AI agent M = 3.65, SE = 0.07, expert M = 4.15, SE = 0.07, friend M = 4.71, SE = 0.07). Ratings were also significantly higher in the performance success condition (M = 5.62, SE = 0.06) than in the failure condition (M = 2.72, SE = 0.06). There was also a significant main effect of trustee characteristics, F(1,594) = 35.13, p <.001, η^2_p = .06, indicating that trustees were rated as warm (M = 4.30, SE = 0.05) rather than competent (M = 4.04, SE = 0.05).

To test hypotheses (5) and (6) – that more warmth would be attributed to humans than AI agents, and more competence to AI agents than humans – we analyzed the interaction between trustee type and trustee characteristics, F(2, 594) = 30.93, p < .001, η^2_p = .09. Comparisons across trustee types showed that AI agents were rated as the least warm, friends as the warmest, and experts in between (p < .001), thus supporting hypothesis (5). AI agents were also rated as less competent than both experts and friends (p < .001), while the competence difference between experts and friends was not significant (p = .27), so our hypothesis (6) was not supported. AI agents and experts were rated as equally warm and competent: for AI agents, warmth (M = 3.65, SE = 0.09) and competence (M = 3.65, SE = 0.08), p = .09; for experts, warmth (M = 4.16, SE = 0.09) and competence (M = 4.14, SE = 0.08), p = .74. In contrast, friends were rated as significantly more warm (M = 5.09, SE = 0.09) than competent (M = 4.33, SE= 0.08), p < .001.

To test hypothesis (7), we analyzed the interaction between trustee type and trustee performance, F(2,594) = 15.81, p < .001, η^2_p = 05. The analysis of simple effects

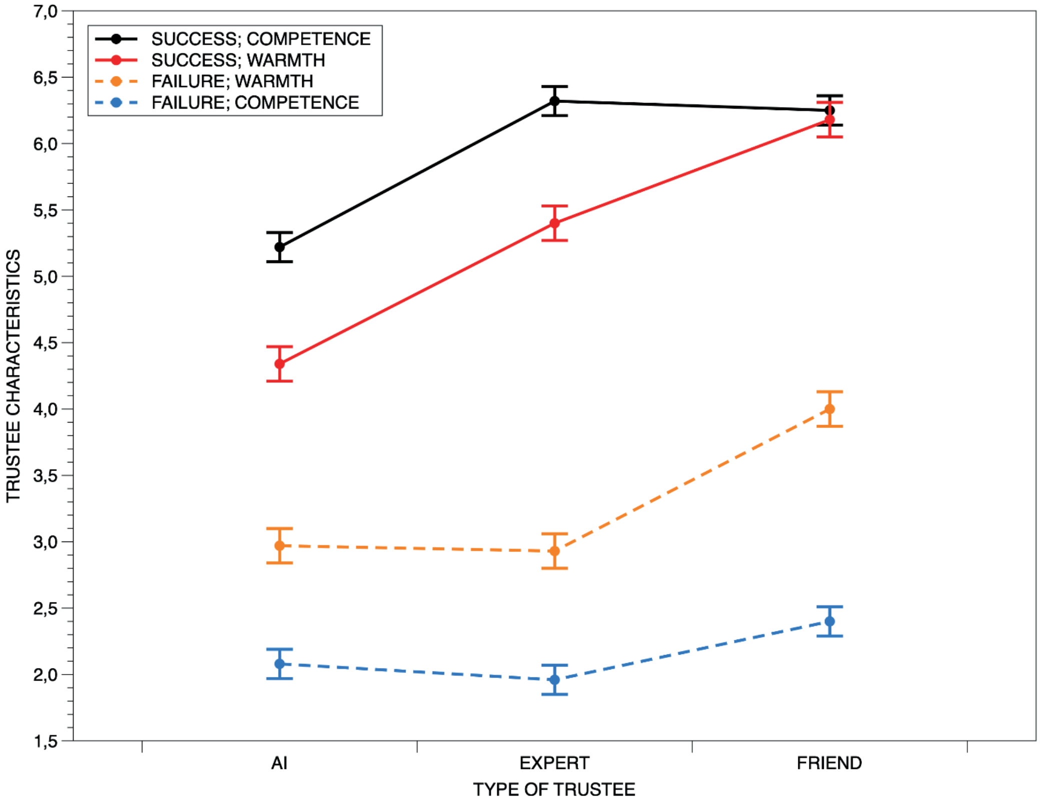

Figure 2. *Warmth (Red and Orange Lines) and Competence (Black and Blue Lines) Ratings for the Three Trustee Types as a Function of Performance (Success: Solid Line; Failure: Dashed Line).*

Note. Error bars represent 95% confidence intervals.

showed that trustee performance significantly influenced trustee characteristic ratings for all three trustee types: ratings were higher in the success condition than in the failure condition for AI agents (M = 4.78, SE = 0.10 vs. M = 2.52, SE = 0.10), $F(1, 594) = 234.97$, $p < .001$, $\eta^2_p = .28$; for experts (M = 5.86, SE = 0.11 vs. M = 2.45, SE = 0.11), $F(1, 594) = 529.39$, $p < .001$, $\eta^2_p = .37$; and for friends (M = 6.21, SE = 0.11 vs. M = 3.20, SE = 0.11), $F(1, 594) = 412.51$, $p < .001$, $\eta^2_p = .41$, but our hypothesis (7) was not supported. Simple effects of trustee type were also significant at both levels of trustee performance. In the success condition, ratings differed between AI agents and experts ($p < .001$) and between AI agents and friends ($p < .001$), but not between experts and friends ($p = .05$), $F(2, 594) = 51.23$, $p < .001$, $\eta^2_p = .15$. In the failure condition, ratings differed between AI agents and friends ($p < .001$) and between experts and friends ($p < .001$), but not between AI agents and experts (p = 1.00), $F(2, 594) = 15.65$, $p < .001$, $\eta^2_p = .05$.

The interaction between trustee characteristics and trustee performance was also significant, $F(1, 594)$ = 394.94, $p < .001$, $\eta^2_p = .40$. Both warmth and competence ratings were higher in the success condition (warmth: M = 5.30, SE = 0.07; competence: M = 5.93, SE = 0.06) than in the failure condition (warmth: M = 3.30, SE = 0.07; competence: M = 2.15, SE = 0.06), $F(1, 594) = 380.12$, $p < .001$, $\eta^2_p = .39$; $F(1, 594) = 1783.57$, $p < .001$, $\eta^2_p = .75$, respectively. Interestingly, in the success condition, trustees were rated as more competent than warm, $F(1, 594) = 97.57$, $p < .001$, $\eta^2_p = .14$, while in the failure condition, trustees were rated as warm rather than competent, $F(1, 594) = 331.71$, $p < .001$, $\eta^2_p = .36$.

The three-way interaction among trustee type, trustee characteristics, and trustee performance was not significant, $F(2, 594) = 0.47$, $p = .63$, $\eta^2_p = .002$.

In summary, all three trustee types received higher ratings on both warmth and competence in the success conditions compared to the failure conditions. In success scenarios, trustees were perceived as more competent than warm, whereas in failure scenarios, they were seen as more warm than competent. Comparing across trustee types, AI agents were rated lower than experts and friends on both dimensions. Notably, friends were rated significantly more warm than competent, while AI agents and experts were perceived as equally warm and competent. We also found that warmth and competence scores are more correlated in the case of friend-related scenarios than in the case of AI agent- or expert-related scenarios.

DISCUSSION

In the present work, we investigated warmth and competence as characteristics of the trustee in trust-related human-AI interactions, as compared to human-human interactions. Across both studies, perceptions of trustee characteristics—warmth and competence—varied systematically by trustee type (human vs. AI) and context. Study 1 showed that while competence was generally rated as more important than warmth, this pattern did not differ significantly between human and AI trustees. Instead, the importance of warmth and competence shifted depending on the situation: warmth mattered most in friend-related scenarios, whereas competence was prioritized in expert- and AI agent-related scenarios. Interestingly, human trustees were rated as warmer than AI agents, and the overall importance assigned to trustee characteristics was slightly higher for humans than for AI agents.

Study 2, focusing on retrospective attributions after trustee performance (success or failure), revealed that all trustee types were rated higher on warmth and competence following success. However, humans (experts and friends) were consistently rated as warmer and more competent than AI agents. Success led to higher competence ratings, while failure increased warmth ratings across trustees, suggesting nuanced attributions depending on outcome. Despite these differences, no three-way interaction among trustee type, characteristics, and performance emerged, indicating similar attribution patterns across trustee types.

The results show that perceptions of a trustee's warmth and competence are not only related, which is consistent with previous evidence (e.g., Abele & Hauke, 2019; Chen & Guo, 2020), but also that their relationship is significantly affected by context. The relationship between warmth and competence was stronger in the case of friend-related situations than in the case of expert- or AI agent related situation, and stronger when assessing trustees retrospectively than in the case of imaginary situations in the present or future. This context-dependency of the halo effect (i.e., the tendency to "think of a person in general as rather good or rather inferior and to color the judgment of the separate qualities by this feeling" Thorndike, 1920, p. 25) in warmth and competence judgments of humans as well as AI agents, is an interesting direction for future research.

Together, these findings suggest that while humans are generally perceived as warmer and more competent than AI agents, contextual factors like the type of situation and performance influence how warmth and competence are perceived when evaluating trustworthiness. The results seem to challenge the notion that warmth is exclusively more important for humans and competence for AI, highlighting a more complex interplay shaped by context and outcomes.

The presented results complement previous research on perceptions of warmth and competence in AI, which found that both dimensions are related to participants' impressions of AI systems (Bergmann et al. 2012; Harris-Watson et al. 2023; Khadpe et al., 2020; McKee et al., 2023) and are important determinants of trust in AI agents (Li et al., 2025). They also support the conclusion of McKee et al. (2023) that the roles AI agents occupy—in our study represented by different trust-related situations—elicit varying impressions of warmth and competence. Our findings further confirm that, within the same context, AI agents are rated lower than humans on both warmth and competence (Ashktorab et al., 2020; Frischknecht, 2021; van der Hoeven, 2018). This suggests persistent human-AI differences in social evaluations, which may affect trust dynamics.

Interestingly, while McKee et al. (2023) found that AI systems tend to be perceived as more competent than warm, our results showed that AI trustees were rated as equally warm and competent, indicating a more balanced perception in trust-related contexts. This balance could reflect growing familiarity with AI capabilities, or the specific nature of trustee roles examined here.

Additionally, the importance of warmth and competence varied significantly by situation type, with warmth prioritized in friend-related contexts and competence in expert- and AI-related contexts. This indicates the contextual and role-dependent nature of trust evaluations, suggesting that trustworthiness assessments are flexible and shaped by the social and task environment.

Our findings might be also interpreted in the context of the functions of stereotypes, which help people make sense of the social world. According to the Stereotype Content Model (Fiske et al., 2023), stereotypes are characterized along two dimensions—warmth and competence—and the combination of these dimensions defines their content. Ingroup and reference groups are typically perceived as both warm and competent, whereas outgroups (e.g., poor people) tend to be seen as low on both dimensions. Groups rated high on competence but low on warmth elicit envious stereotypes (e.g., professionals), while those rated high on warmth but low on competence evoke paternalistic stereotypes (e.g., children, elderly, housewives).

In our research, AI agents were perceived as less competent than humans from typically high-competence groups (friends or experts), less warm than friends (a high-warmth group), and equally or less warm than experts (a lower-warmth group). As AI technologies are increasingly viewed as social actors, with people interacting with them in a similar manner as with humans, it is valuable to further investigate the stereotypes associated with AI agents as a distinct social group.

The results of our present research may also be understood through the lens of cognitive biases. Recent studies show that the label "AI" can carry negative connotations in various contexts. For example, participants reported lower trust in emails authored with AI assistance (Liu et al., 2022) and expressed more favorable judgments of human-created versus AI-created artwork (Bellaiche et al., 2023; Millet et al., 2023). Interestingly, even when AI-generated artworks were rated more positively overall, participants rated pieces they believed to be AI-created more negatively than those perceived as human-made (Grassini & Koivisto, 2024). This "anti-AI bias" may stem from a dissociation between explicit and implicit attitudes toward AI (Fietta et al., 2021): although over 70% of participants explicitly expressed positive attitudes toward AI, 85% demonstrated implicit negative or neutral biases, suggesting that attitudes toward AI are influenced by unconscious biases. Our findings align with this pattern, as AI agents were rated lower than humans on both warmth and competence under identical conditions, supporting the presence of an implicit anti-AI bias in impression formation. Moreover, the reduced emphasis on warmth and competence in trust-related evaluations of AI agents may reflect a broader apprehension toward AI technologies. Future research should explore whether these negative attitudes and diminished trust in AI, relative to humans, are rooted in unconscious biases and how such biases can be mitigated to foster more equitable human-AI interactions.

The main limitation of the present research is that it did not involve real interactions with AI agents, thus limiting its ecological validity. In the presented studies participants did not interact with other people or AI agents, but were instructed to imagine such interactions, so the measures were declaratory, and some terms and scenarios were not precisely specified. The generality of the term "AI agent" itself introduces variability, as participants might have envisioned very different technologies based on their personal experiences and knowledge, potentially activating diverse associations. Extending this line of research to real interactions is especially important in the context of algorithm aversion (Dietvorst et al., 2015, 2016). People often choose to interact with humans over AI agents, despite knowing that algorithms can carry out the task better, and even more so as the stakes in the interaction are higher, which is referred to as the tragedy of algorithm aversion (Filiz et al., 2023). This suggests that in case of human-AI interactions declarative (vs. behavioral) measures could lead to erroneous conclusions, especially in the case of high-stakes situations. Future research should aim to overcome these limitations by involving real human – AI interactions, employing behavioral measures and incorporating real-life stakes.

Another limitation of the present research is that we focused solely on one aspect of warmth—sociability—which was a natural choice given the AI context. However, warmth also comprises a second important dimension: morality (Brambilla et al., 2011; Leach et al., 2007), which may be equally influential in impression formation (Goodwin et al., 2014). Future studies incorporating the morality aspect of warmth could provide deeper insights into the fundamental dimensions of social cognition relevant to trust in human-AI interactions.

Future research could also benefit from fine-tuning the design of the experimental materials. While the vignettes used in the presented studies were designed to be realistic, relatable, and including a broad spectrum of situations requiring trust, we did not control many of their other aspects. It would be beneficial to control, for instance, the high or low stakes of analyzed situations (Filiz et al., 2023), a gains or losses framing (Kahneman & Tversky, 1979), and the presence of market sociality cues that could trigger a market mindset (Gasiorowska & Zaleskiewicz, 2021; Zaleskiewicz et al., 2020, 2022).

Finally, the present research was conducted primarily among native English speakers from the US and UK, which are typically individualistic cultures. Since individualistic cultures tend to emphasize competence, while collectivistic cultures prioritize warmth (Hofstede, 1980; Wojciszke, 1997), it remains unclear whether our findings generalize to collectivistic contexts. Future cross-cultural research is needed to validate and extend these results.

CONCLUSION

Understanding the factors that affect people's perceptions of AI technologies is essential for fostering effective and harmonious human-AI interactions in increasingly AI-integrated societies. In this research, we examined trust-related interactions involving both human and AI trustees, focusing on warmth and competence—the two fundamental dimensions of social cognition known to guide impression formation and trust.

Our results demonstrate that warmth and competence play a less prominent role in shaping trust judgments during human-AI interactions compared to human-human interactions. Moreover, the relative importance of these dimensions varies depending on the situational context, suggesting that people flexibly adjust their trust criteria based on the nature of the interaction. While AI agents are generally perceived as equally warm and competent when considered in isolation, they are consistently rated lower on both dimensions compared to human trustees. This indicates a persistent trust gap between AI and humans, highlighting challenges in achieving parity in social evaluations of AI systems.

Additionally, perceptions of AI trustees' warmth and competence appear less sensitive to situational outcomes (success or failure) than those of human trustees, possibly reflecting more rigid or implicit biases toward AI. This diminished responsiveness could signal a form of apprehension or skepticism toward AI, with implications for the design and deployment of AI systems in social contexts.

By integrating insights from social psychology with AI research, these findings contribute to a deeper understanding of the psychological determinants of trust in AI. Such knowledge may be critical for developing AI technologies that are not only effective but also socially accepted and trusted by human users.

COMPLIANCE WITH ETHICAL STANDARDS

This research was carried out in accordance with the recommendations of the Declaration of Helsinki and its later amendments with written informed consent from all participants. All participants were adults. The protocols were approved by the Ethics Committee for Research at the Faculty of Psychology in Wroclaw of SWPS University, decision number 04/P/03/2024.

DATA AVAILABILITY

The data underlying presented results presented are openly available in Figshare at https:/doi.org/10.6084/m9.figshare.26139574.

REFERENCES

Abele, A. E., & Hauke, N. (2019). Comparing the facets of the big two in global evaluation of self versus other people. *European Journal of Social Psychology*, *50*(5), 969–982. https://doi.org/10.1002/ejsp.2639

Ashktorab, Z., Liao, Q. V., Dugan, C., Johnson, J., Pan, Q., Zhang, W., Kumaravel, S., & Campbell, M. (2020). Human-AI collaboration in a cooperative game setting. *Proceedings of the ACM on Human-Computer Interaction, 4*(CSCW2), 1–20. https://doi.org/10.1145/3415167

Bellaiche, L., Shahi, R., Turpin, M. H., Ragnhildstveit, A., Sprockett, S., Barr, N., Christensen, A., & Seli, P. (2023). Humans versus AI: Whether and why we prefer human-created compared to AI-created artwork. *Cognitive Research: Principles and Implications, 8*(1). https://doi.org/10.1186/s41235-023-00499-6

Bergmann, K., Eyssel, F., & Kopp, S. (2012). A second chance to make a first impression? How appearance and nonverbal behavior affect perceived warmth and competence of virtual agents over time. In Y. Nakano, M. Neff, A. Paiva, & M. Walker (Eds.), *Intelligent virtual agents* (Vol. 7502, pp. 126–138). Springer. https://doi.org/10.1007/978-3-642-33197-8_13

Blanca, M., Alarcón, R., Arnau, J., Bendayan, R., & Bono, R. (2017). Non-normal data: Is ANOVA still a valid option? *Psicothema*, *29*(4), 552–557. https://doi.org/10.7334/psicothema2016.383

Brambilla, M., Rusconi, P., Sacchi, S., & Cherubini, P. (2011). Looking for honesty: The primary role of morality (vs. sociability and competence) in information gathering. *European Journal of Social Psychology, 41*(2), 135–143. https://doi.org/10.1002/ejsp.744

Camerer, C. F., Dreber, A., Holzmeister, F., Ho, T.-H., Huber, J., Johannesson, M., Kirchler, M., Nave, G., Nosek, B. A., Pfeiffer, T., Altmejd, A., Buttrick, N., Chan, T., Chen, Y., Forsell, E., Gampa, A., Heikensten, E., Hummer, L., Imai, T., & Isaksson, S. (2018). Evaluating the replicability of social science experiments in Nature and Science between 2010 and 2015. *Nature Human Behaviour*, *2*(9), 637–644. https://doi.org/10.1038/s41562-018-0399-z

Chen, F., & Guo, T. (2020). Effects of competence information on perceptions of warmth. *Asian Journal of Social Psychology*, *24*(4). https://doi.org/10.1111/ajsp.12452

Christoforakos, L., Gallucci, A., Surmava-Große, T., Ullrich, D., & Diefenbach, S. (2021). Can robots earn our trust the same way humans do? A systematic exploration of competence, warmth, and anthropomorphism as determinants of trust development in HRI. *Frontiers in Robotics and AI, 8*. https://doi.org/10.3389/frobt. 2021.640444

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Lawrence Erlbaum Associates.

Dietvorst, B. J., Simmons, J. P., & Massey, C. (2015). Algorithm aversion: People erroneously avoid algorithms after seeing them err. *Journal of Experimental Psychology: General*, *144*(1), 114–126. https://doi.org/10.1037/xge0000033

Dietvorst, B. J., Simmons, J. P., & Massey, C. (2016). Overcoming Algorithm Aversion: People Will Use Imperfect Algorithms If They Can (Even Slightly) Modify Them. *Management Science*, *64*(3), 1155–1170. https://doi.org/10.1287/mnsc.2016.2643

Espina-Romero, L., Noroño Sánchez, J. G., Gutiérrez Hurtado, H., Dworaczek Conde, H., Solier Castro, Y., Cervera Cajo, L. E., & Rio Corredoira, J. (2023). Which industrial sectors are affected by artificial intelligence? A bibliometric analysis of trends and perspectives. *Sustainability, 15*(16), 12176. https://doi.org/10.3390/su151612176

Fietta, V., Zecchinato, F., Stasi, B. D., Polato, M., & Monaro, M. (2021). Dissociation between users' explicit and implicit attitudes toward artificial intelligence: An experimental study. *IEEE Transactions on Human-Machine Systems, 52*(3), 1–9. https://doi.org/10.1109/thms.2021.3125280

Filiz, I., Judek, J. R., Lorenz, M., & Spiwoks, M. (2023). The extent of algorithm aversion in decision-making situations with varying gravity. *PLOS ONE*, *18*(2), e0278751–e0278751. https://doi.org/10.1371/journal.pone.0278751

Fiske, S. (1993). Social cognition and social perception. *Annual Review of Psychology, 44*(1), 155–194. https://doi.org/10.1146/annurev.psych.44.1.155

Fiske, S. T. (2018). Stereotype content: Warmth and competence endure. *Current Directions in Psychological Science, 27*(2), 67–73. https://doi.org/10.1177/0963721417738825

Fiske, S. T., Cuddy, A. J. C., & Glick, P. (2007). Universal dimensions of social cognition: Warmth and competence. *Trends in Cognitive Sciences, 11*(2), 77–83. https://doi.org/10.1016/j.tics.2006.11.005

Fiske, S. T., Cuddy, A. J. C., Glick, P., & Xu, J. (2002). A model of (often mixed) stereotype content: Competence and warmth respectively follow from perceived status and competition. *Journal of Personality and Social Psychology, 82*(6), 878–902. https://doi.org/10.1037//0022-3514.82.6.878

Frischknecht, R. (2021). A social cognition perspective on autonomous technology. *Computers in Human Behavior, 122,* 106815. https://doi.org/10.1016/j.chb.2021.106815

Gasiorowska, A., & Zaleskiewicz, T. (2021). Trading in search of structure: Market relationships as a compensatory control tool. *Journal of Personality and Social Psychology*, *120*(2), 300–334. https://doi.org/10.1037/pspi0000246

Gilad, Z., Amir, O., & Levontin, L. (2021). The effects of warmth and competence perceptions on users' choice of an AI system. *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems. ACM.* https://doi.org/10.1145/3411764.3446863

Goodwin, G. P., Piazza, J., & Rozin, P. (2014). Moral character predominates in person perception and evaluation. *Journal of Personality and Social Psychology, 106*(1), 148–168. https://doi.org/10.1037/a0034726

Grassini, S., & Koivisto, M. (2024). Understanding how personality traits, experiences, and attitudes shape negative bias toward AI-generated artworks. *Scientific Reports, 14*(1), 4113. https://doi.org/10.1038/s41598-024-54294-4

Hancock, P. A., Billings, D. R., Schaefer, K. E., Chen, J. Y. C., de Visser, E. J., & Parasuraman, R. (2011). A meta-analysis of factors affecting trust in human-robot interaction. *Human Factors, 53*(5), 517–527. https://doi.org/10.1177/0018720811417254

Harris-Watson, A. M., Larson, L. E., Lauharatanahirun, N., DeChurch, L. A., & Contractor, N. S. (2023). Social perception in human-AI teams: Warmth and competence predict receptivity to AI teammates. *Computers in Human Behavior, 145*, 107765. https://doi.org/10.1016/j.chb.2023.107765

Hofstede, G. (1980). *Culture's consequences: International differences in work-related values.* Sage Publications.

Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica, 47*(2), 263–292.

Khadpe, P., Krishna, R., Fei-Fei, L., Hancock, J. T., & Bernstein, M. S. (2020). Conceptual metaphors impact perceptions of human-AI collaboration. *Proceedings of the ACM on Human-Computer Interaction, 4*(CSCW2), 1–26. https://doi.org/10.1145/3415234

Kulms, P., & Kopp, S. (2018). A social cognition perspective on human–computer trust: The effect of perceived warmth and competence on trust in decision-making with computers. *Frontiers in Digital Humanities, 5*. https://doi.org/10.3389/fdigh.2018.00014

Leach, C. W., Ellemers, N., & Barreto, M. (2007). Group virtue: The importance of morality (vs. competence and sociability) in the positive evaluation of in-groups. *Journal of Personality and Social Psychology, 93*(2), 234–249. https://doi.org/10.1037/0022-3514.93.2.234

Li, Y., Wu, B., Huang, Y., Liu, J., Wu, J., & Luan, S. (2025). Warmth, competence, and the determinants of trust in artificial intelligence: A cross-sectional survey from China. *International Journal of Human–Computer Interaction, 41*(8), 5024–5038. https://doi.org/10.1080/10447318.2024.2356909

Liu, Y., Mittal, A., Yang, D., & Bruckman, A. (2022). Will AI console me when I lose my pet? Understanding perceptions of AI-mediated email writing. *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems.* ACM. https://doi.org/10.1145/3491102.3517731

Lovakov, A., & Agadullina, E. R. (2021). Empirically derived guidelines for effect size interpretation in social psychology. *European Journal of Social Psychology*, *51*(3), 485–504. https://doi.org/10.1002/ejsp.2752

Malik, S., Muhammad, K., & Waheed, Y. (2024). Artificial intelligence and industrial applications: A revolution in modern industries. *Ain Shams Engineering Journal, 15*(9), 102886. https://doi.org/10.1016/j.asej.2024.102886

McKee, K. R., Bai, X., & Fiske, S. T. (2023). Humans perceive warmth and competence in artificial intelligence. *iScience, 26*(8), 107256. https://doi.org/10.1016/j.isci.2023.107256

Millet, K., Buehler, F., Du, G., & Kokkoris, M. (2023). Defending humankind: Anthropocentric bias in the appreciation of AI art. *Computers in Human Behavior, 143,* 107707. https://doi.org/10.1016/j.chb.2023.107707

Nass, C., Steuer, J., & Tauber, E. R. (1994). Computers are social actors. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems: Celebrating Interdependence* (pp. 72–78). ACM. https://doi.org/10.1145/191666.191703

Osborne, J. (2008). *Best Practices in Quantitative Methods* (2nd ed.). SAGE.

Reeves, B., & Nass, C. I. (1996). *The media equation: How people treat computers, television, and new media like real people and places.* CSLI Publications.

Richard, F. D., Bond, C. F., & Stokes-Zoota, J. J. (2003). One Hundred Years of Social Psychology Quantitatively Described. *Review of General Psychology*, *7*(4), 331–363. https://doi.org/10.1037/1089-2680.7.4.331

Thorndike, E. L. (1920). A constant error in psychological ratings. *Journal of Applied Psychology*, *4*(1), 25–29. https://doi.org/10.1037/h0071663

van der Hoeven, B. (2018). *One hamburger with extra cogwheels, please: The difference in warmth and competence ratings between humans and artificial intelligence in the workplace* [Unpublished doctoral dissertation]. Tilburg University.

Wojciszke, B. (1997). Parallels between competence- versus morality-related traits and individualistic versus collectivistic values. *European Journal of Social Psychology, 27*(3), 245–256. https://doi.org/10.1002/(sici)1099-0992(199705)27:3<245::aid-ejsp819>3.0.co;2-h

Zaleskiewicz, T., Gasiorowska, A., & Kuzminska, A. (2022). Market mindset reduces endorsement of individualizing moral foundations, but not in liberals. *Journal of Social and Political Psychology*, *10*(2), 743–759. https://doi.org/10.5964/jspp.8163

Zaleskiewicz, T., Gasiorowska, A., Kuzminska, A. O., Korotusz, P., & Tomczak, P. (2020). Market mindset impacts moral decisions: The exposure to market relationships makes moral choices more utilitarian by means of proportional thinking. European Journal of Social Psychology. https://doi.org/10.1002/ejsp.2701

APPENDIX

Pilot Study

Participant instructions

Please imagine yourself in each of these situations. Whose advice would you rather follow?
Group 1 (N = 100): (a) Human or (b) AI agent
Group 2 (N = 100): (a) Friend or (b) Expert

Results

Vignette	HUMAN (VS AI)	FRIEND (VS EXPERT)	TYPE*
You've come across a technical, work-related issue and you're looking for a way to address it.	40%	4%	AI
You received a significant amount of extra money and you don't know how to invest it.	41%	7%	AI
You have an impossible deadline at work and you need help completing the project on time.	36%	21%	AI
You need to learn a new language fast and you don't know how to approach it.	49%	6%	AI
You want to learn a new skill and you need advice on how to do it.	44%	15%	AI
You have a legal problem that could lead to serious consequences and you don't know how to handle it.	79%	7%	E
You have a non-life-threatening health issue and you don't know what to do about it.	77%	15%	E
You have a long-term conflict with a colleague and you don't know how to handle it.	94%	45%	E
You have a difficult financial situation and you're looking for ways out of it.	59%	23%	E
You need to improve your personal safety and you need to choose the right measures.	66%	22%	E
You take care of your family member and you need someone to take over your duties while you're away for a couple of days.	94%	64%	F
You had a fight with a close friend and you need advice regarding how to handle it.	93%	83%	F
You're at a crossroads in your important long-term relationship and you need advice about it.	94%	56%	F
You have a free weekend and want to use it to increase your wellbeing and you don't know how to spend it.	53%	75%	F
You like your job, but you got an alternative offer that you're considering, you can't make up your mind.	80%	60%	F

Note. *TYPE: AI = AI agent preference; E = expert preference; F = friend preference. Percentages in cells refer to percentage of (a) answers, i.e., "Human" in Group 1, and "Friend" in Group 2. Underlined vignettes are the ones used in present research.

Study 1 Normality Tests

	K-S test (df = 198)	Skewness (SE)	Kurtosis (SE)
WARMTH TYPE: AI	0.10 ***	-0.35(0.17)	-0.60(0.34)
WARMTH TYPE: E	0.12 ***	-0.52(0.17)	-0.47(0.34)
WARMTH TYPE: F	0.18 ***	-1.02(0.17)	0.25(0.34)
COMPETENCE TYPE: AI	0.24 ***	-1.73(0.17)	2.43(0.34)
COMPETENCE TYPE: E	0.25 ***	-1.64(0.17)	2.07(0.34)
COMPETENCE TYPE: F	0.17 ***	-1.06(0.17)	0.33(0.34)

Note. *** $p < .001$; TYPE: AI = AI agent preference; E = expert preference; F = friend preference.

Study 2 Normality Tests

	K-S test (df = 600)	Skewness (SE)	Kurtosis (SE)
WARMTH	0.06 ***	-0.23(0.10)	-0.89(0.20)
COMPETENCE	0.14 ***	-1.03(0.10)	-1.58(0.20)

Note. *** $p < .001$;