Mind(sets) Over Machine?

The Influence of Implicit Self-Theories in Human-Robot Interaction

By Dwain D. Allan

UC UNIVERSITY OF CANTERBURY
Te Whare Wānanga o Waitaha
CHRISTCHURCH NEW ZEALAND

A thesis submitted in partial fulfilment of the requirements for the degree of Doctor of Philosophy at the Department of Psychology, Faculty of Science, University of Canterbury, Christchurch, Aotearoa New Zealand
June 2022

Mind(sets) Over Machine? The Influence of Implicit Self-Theories in Human-Robot Interaction

By
Dwain D. Allan

A thesis submitted in partial fulfillment of the requirements
for the degree of
Doctor of Philosophy

Department of Psychology
Faculty of Science
University of Canterbury
June 2022

UC
UNIVERSITY OF CANTERBURY
Te Whare Wānanga o Waitaha
CHRISTCHURCH NEW ZEALAND

Abstract

Implicit self-theory asserts that an individual's underlying beliefs about whether self-attributes (e.g., personality and intelligence) are fixed (entity theory) or mutable (incremental theory) causally affect motivation and behavior—with the most profound effects emerging in situations that involve challenges and setbacks. In support of this notion, several lines of research suggest that these beliefs hold some influence over people's perception and behavior in diverse domains such as education, brand acceptance, and financial decision-making, among others. It is, however, presently unknown whether implicit self-theories exert such influence on people's experiences of social robots. To address this gap, this research tested, in a series of three studies, the proposition that implicit self-theories represent an important variable, that influences the manner in which one perceives and responds to social robots. Study 1 provided the first evidence that an individual's implicit self-theory orientation influences their perception of emerging social robots developed for everyday use. In particular, those endorsing more of an entity theory expressed greater robot anxiety than those endorsing more of an incremental theory. This finding held even when controlling for a range of covariate influences. In addition, incremental theorists, compared to entity theorists responded more favorably to social robots in general. Study 2 built on and substantively extended the findings of Study 1 by examining the effects of implicit self-theories on people's responses to a robot that praised them for ability (i.e., intelligence), or for effort (i.e., hard work), after completing a difficult task. Results revealed that entity theorists evaluated a robot that delivered ability praise as more likable and intelligent than one that delivered effort praise. However, incremental theorists were unaffected by either praise type and rated the robot favorably regardless of the praise it delivered. Study 3, expanded the findings of Studies 1 and 2 to investigate the impact of implicit self-theories on people's responses to a robot that defeats human beings in a general knowledge quiz game. Results showed that incremental theorists, compared to entity theorists were more likely to indicate an interest in playing against the robot after imagining losing to it. Whereas entity theorists rated such robots as presenting more identity and realistic threats. Together, these studies extend and enrich the Human-Robot Interaction (HRI) literature by establishing implicit self-theories as an important and meaningful variable for which to advance the understanding of HRI today. In so doing, this research attempts to respond to the ever-increasing demand for research on the psychological variables that underlie how people perceive and interact with robots—which, in many ways, has special urgency given the inexorable rise of AI and robotics in the social domain of everyday experience. In consequence, findings may contribute to the design of new or improved social robots that can reflect or shape beliefs, and, hence, build a greater sense of identification and trust with the intended human user.

Keywords: implicit self-theories, mindset, human–robot interaction, social robots, robot acceptance, robot anxiety, praise, effort-ability, identity threat, realistic threat, perception

UC UNIVERSITY OF CANTERBURY
Te Whare Wānanga o Waitaha
CHRISTCHURCH NEW ZEALAND

Deputy Vice-Chancellor's Office
Postgraduate Office

Co-Authorship Form

This form is to accompany the submission of any thesis that contains research reported in co-authored work that has been published, accepted for publication, or submitted for publication. A copy of this form should be included for each co-authored work that is included in the thesis. Completed forms should be included at the front (after the thesis abstract) of each copy of the thesis submitted for examination and library deposit.

Please indicate the chapter/section/pages of this thesis that are extracted from co-authored work and provide details of the publication or submission from the extract comes:

Primarily Chapter 2, in addition to parts of Chapters 1 and 5.

Allan, D.D., Vonasch, A.J. & Bartneck, C. The Doors of Social Robot Perception: The Influence of Implicit Self-theories. *International Journal of Social Robotics* 14, 127–140 (2022). doi: 10.1007/s12369-021-00767-9

Please detail the nature and extent (%) of contribution by the candidate:
The candidate was the primary driver of every aspect of this work. 90%

Certification by Co-authors:
If there is more than one co-author then a single co-author can sign on behalf of all
The undersigned certifys that:

- The above statement correctly reflects the nature and extent of the PhD candidate's contribution to this co-authored work
- In cases where the candidate was the lead author of the co-authored work he or she wrote the text

Andy Vonasch

Name: Andrew Vonasch **Signature:** **Date:** 22/3/2022

UNIVERSITY OF CANTERBURY
Te Whare Wānanga o Waitaha
CHRISTCHURCH NEW ZEALAND

Deputy Vice-Chancellor's Office
Postgraduate Office

Co-Authorship Form

This form is to accompany the submission of any thesis that contains research reported in co-authored work that has been published, accepted for publication, or submitted for publication. A copy of this form should be included for each co-authored work that is included in the thesis. Completed forms should be included at the front (after the thesis abstract) of each copy of the thesis submitted for examination and library deposit.

Please indicate the chapter/section/pages of this thesis that are extracted from co-authored work and provide details of the publication or submission from the extract comes:

Primarily Chapter 3, in addition to parts of Chapters 1 and 5.

Allan, D. D., Vonasch, A. J., & Bartneck, C. (2022). "I Have to Praise You Like I Should?" The Effects of Implicit Self-Theories and Robot-Delivered Praise on Evaluations of a Social Robot. *International Journal of Social Robotics*, 1-12. doi: 10.5555/3523760.3523792

Please detail the nature and extent (%) of contribution by the candidate:
The candidate was the primary driver of every aspect of this work. 90%

Certification by Co-authors:
If there is more than one co-author then a single co-author can sign on behalf of all
The undersigned certifys that:

- The above statement correctly reflects the nature and extent of the PhD candidate's contribution to this co-authored work
- In cases where the candidate was the lead author of the co-authored work he or she wrote the text

Andy Vonasch

Name: Andrew Vonasch **Signature:** **Date:** 22/3/2022

Deputy Vice-Chancellor's Office
Postgraduate Office

Co-Authorship Form

This form is to accompany the submission of any thesis that contains research reported in co-authored work that has been published, accepted for publication, or submitted for publication. A copy of this form should be included for each co-authored work that is included in the thesis. Completed forms should be included at the front (after the thesis abstract) of each copy of the thesis submitted for examination and library deposit.

Please indicate the chapter/section/pages of this thesis that are extracted from co-authored work and provide details of the publication or submission from the extract comes:

Primarily Chapter 4, in addition to parts of Chapters 1 and 5.

Allan, D. D., Vonasch, A. J., & Bartneck, C. (2022, March). Better than Us: The Role of Implicit Self-Theories in Determining Perceived Threat Responses in HRI. In *Proceedings of the 2022 ACM/IEEE International Conference on Human-Robot Interaction* (pp. 215-224). doi: 10.1007/s12369-021-00848-9

Please detail the nature and extent (%) of contribution by the candidate:
The candidate was the primary driver of every aspect of this work. 90%

Certification by Co-authors:
If there is more than one co-author then a single co-author can sign on behalf of all
The undersigned certifys that:

- The above statement correctly reflects the nature and extent of the PhD candidate's contribution to this co-authored work
- In cases where the candidate was the lead author of the co-authored work he or she wrote the text

Andy Vonasch

Name: Andrew Vonasch **Signature:** **Date:** 22/3/2022

Declaration

I hereby declare that this thesis is my own work and has not been submitted in whole or in part for consideration for any other degree or qualification in this, or any other institution.

Dwain D. Allan
June 2022

Acknowledgements

I wish to thank my supervisors, Christoph Bartneck and Andy Vonasch for supporting this research and for aiding me in my development as a scientist. It is also necessary to thank my friends and family for their encouragement and good wishes during this undertaking. Recognition and gratitude is due, finally, and most significantly, to my wife Alana, who, in my contention, is an absolute paragon of strength, generosity, care and unconditional support—in all respects—and who, in truth, cannot be thanked enough.

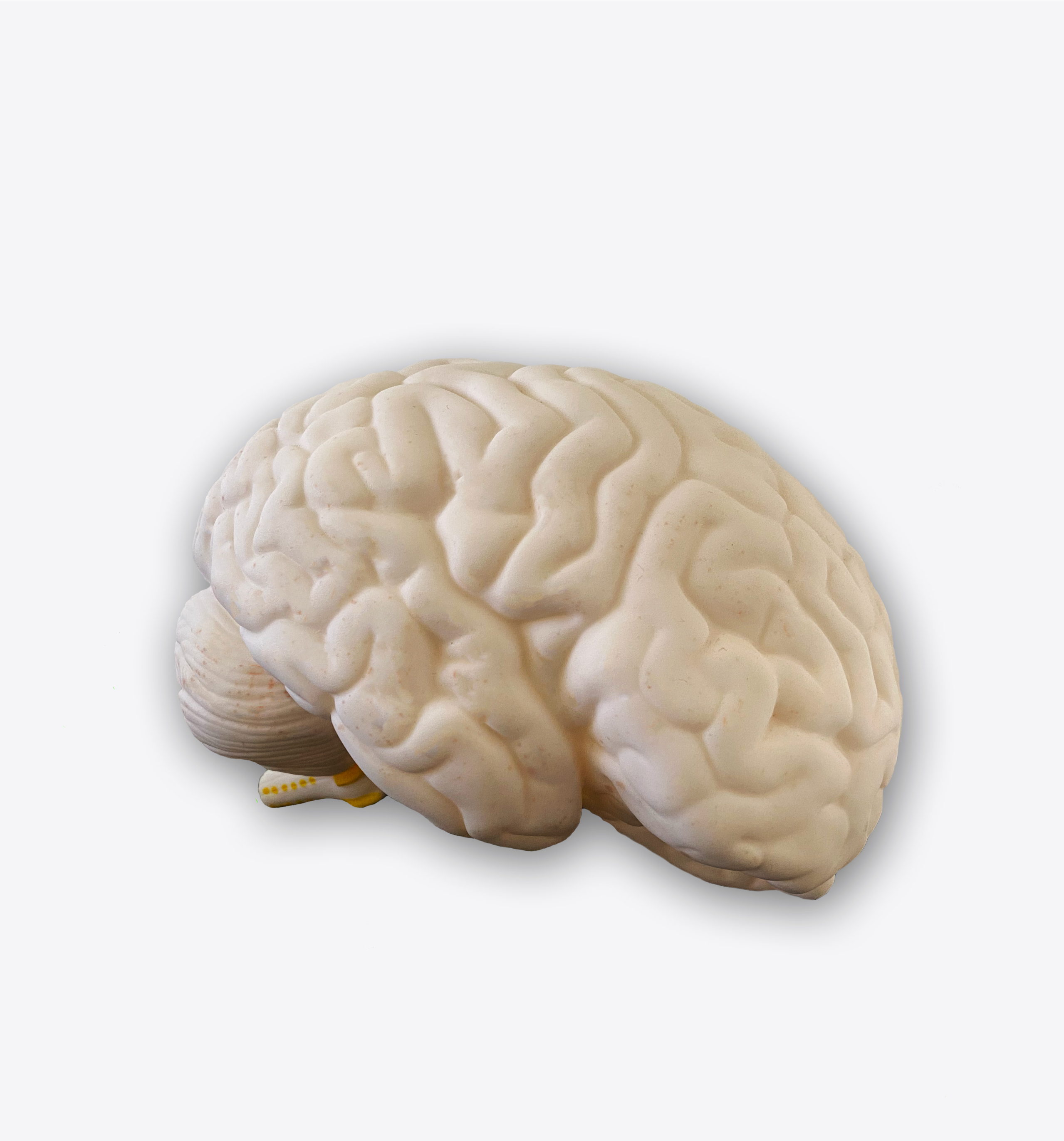
Contents

List of Tables

List of Figures

Chapter I

Introduction



I

Introduction

1.1 General Introduction

Understanding people's impressions and evaluations of social robots[1] and thus, their responses toward them, is a central theme of continuing and recently revived importance in the field of Human-Robot Interaction (henceforth, HRI). (For examples see Dautenhahn, 2007; Belpaeme et al., 2020; Naneva et al., 2020; Spatola and Wudarczyk, 2021; Manzi et al., 2021; Zhong et al., 2022). Recent research suggests that this topic has special urgency given the rise of robotic technologies (e.g., Haegele, 2016; Beraldo et al., 2019; de Graaf et al., 2019; Morsunbul, 2019; Ghazali et al., 2020; Brondi et al., 2021a; Stock-Homburg et al., 2022) and public debate surrounding the appropriateness of automation (e.g., cars, drones, robots) in society and culture (see Horowitz, 2016; Yogeeswaran et al., 2016; Floridi, 2017; Złotowski et al., 2017; Galin and Mamchenko, 2021; Dang and Liu, 2021; van Wynsberghe et al., 2022). This in turn, has lead to a reasonable degree of consensus regarding the pressing need for research to advance understanding of how people view and respond to social robots (see Andrist et al., 2015; Reich-Stiebert and Eyssel, 2015; Ferrari et al., 2016; Appel et al., 2020; Bartneck et al., 2020; Hinks, 2020; Smith et al., 2021; Leoste et al., 2021; Schepers and Streukens, 2022).

Cumulative research thus far has established that certain individual difference variables[2](e.g., prior robot experience, personality dispositions, and demographic characteristics) strongly affect people's thoughts and behaviors regarding social robots (e.g., Eurobarometer, 2012; Kuo et al., 2009; Bartneck et al., 2007; Mutlu et al., 2006; Morsunbul, 2019; Esterwood and Robert, 2022).

To be sure, identifying and investigating such psycho-social variables is of fundamental import to the field. Why? Most basically: these factors directly or indirectly affect acceptance and usage of social robots (Latikka et al., 2019; Esterwood et al., 2021; Forgas-Coll et al., 2022). This is arguably the center of HRI research today (see for example Brondi et al., 2021b; Nomura and Tanaka, 2022; Lin et al., 2021;

[1] Although there are numerous definitions for the term social robot (e.g., Bartneck and Forlizzi, 2004; Billard and Dautenhahn, 1997; Duffy, 2003; Breazeal, 2003; Lee et al., 2006; Zhao, 2006), a social robot can generally be described as a physically embodied, autonomous or semi-autonomous agent designed (mostly) to interact socially with human beings. Note also, the terms "robot" and "social robot" are used here, interchangeably.

[2] Incidentally, some have argued (e.g., Collins, 2019; Xu, 2019; Matthews et al., 2020), in particular, that individual difference research in HRI is, at present, only partly understood and warrants further investigation.

Saari et al., 2022; He et al., 2022), hence the motivation for conducting this research.

Now with that established, there are good reasons to suspect that implicit self-theories—underlying beliefs about whether self-attributes are fixed or are malleable—may play an influential role in shaping perceptions of, and responses to, social robots.

Social psychology research has repeatedly shown, for example, that implicit self-theories affect people's information processing (e.g., Poon and Koehler, 2006), attitude formation (e.g., Dweck, 1986), and social judgments towards themselves and others (e.g., Dweck et al., 1995a). Crucially, however, the precise role of implicit self-theories in HRI has yet to be determined.

That said, two studies, have recently assessed the use of robots to help foster a growth mindset (known more formally as an incremental theory, described below in Section 1.2) in children with respect to learning (Park et al., 2017; Davison et al., 2021). As discussed later in this chapter, these studies did not examine the influence of implicit self-theories on how people view and respond to robots. Consequently, it remains unclear whether implicit self-theories impact these factors.

In response to this gap, this thesis investigates, in a series of three studies, the proposition that implicit self-theory is an influential variable in determining individuals' perceptions, evaluations, and responses toward social robots. In doing so, this research seeks to provide a number of theoretical and practical implications relevant to the field of HRI.

The rest of this chapter is organized as follows: Section 1.2 presents an overview of implicit self-theories and subsequently builds the theoretical background regarding the potential influence of implicit self-theories in HRI. Lastly, Section 1.3 provides an overview of the empirical studies conducted to assess the central proposition of this thesis.

1.2 Theoretical Background

Implicit Self-Theories

Research has consistently demonstrated that people hold differing beliefs about the fixed or mutable nature of an array of human traits and self-attributes such as intelligence (Blackwell et al., 2007; Dweck et al., 1995b; Robins and Pals, 2002), personality (Erdley et al., 1997; Chiu et al., 1997b), morality (Chiu et al., 1997a; Huang et al., 2017), emotions (King and dela Rosa, 2019; Tamir et al., 2007; De Castella et al., 2013), and relationships (Knee, 1998; Knee et al., 2003; Ng and Tong, 2013), among others. These types of beliefs are referred to as implicit self-theories (Dweck and Leggett, 1988; Molden and Dweck, 2006)—or, more colloquially, mindsets.[3] They

[3]The terms "implicit self-theories" or "implicit self-theory" are used herein, rather than "mindset," as the latter is an ambiguous term with numerous conceptualizations and diverse meanings (see for example Freitas et al., 2004; Gollwitzer, 2012; Murphy and Dweck, 2016a; Rucker and Galinsky, 2016). Accordingly, the terms "incremental theories" and "entity theories," are used here, instead of "fixed" and "growth" mindsets (see also Wheeler and Omair, 2016).

are described as implicit because most individuals are unaware of them [4] (e.g., they are not evaluative, in the way that, say, attitudes are; Plaks et al., 2005), and as theories, since they are in essence, falsifiable ideas about what a specific trait or attribute is and how it might work (Dweck and Yeager, 2019). Broadly speaking, implicit self-theories exist on a spectrum from the incremental theory (i.e., incremental theorists), which assumes that self-attributes are responsive to change through concerted effort and education, to the entity theory (i.e., entity theorists), which assumes that self-attributes are largely fixed and immutable (for review see Dweck, 2017; Dweck and Yeager, 2019).

More than 40 years of correlational and experimental studies provide evidence for the validity of implicit self-theories (e.g., Dweck et al., 1995b, 1993; Dweck and Leggett, 1988; Dweck and Yeager, 2019). Indeed, evidence indicates that most individuals will hold both self-theories concurrently (Yorkston et al., 2010), but will chronically endorse one self-theory or the other (Dweck et al., 1995b; Murphy and Dweck, 2016b), the consequences of which will arise, typically outside of one's direct awareness (Rucker and He, 2016). Moreover, the occurrence of incremental and entity self-theories in a given population appears to be roughly equal (Burnette et al., 2013).

As well, researchers (e.g., Murphy and Dweck, 2016a) have established that implicit self-theories are distinct from seemingly similar, well-established theories and constructs such as the Big Five Model of personality (e.g., Spinath et al., 2003), regulatory focus theory, and dual process models (e.g., Mathur et al., 2016), self-efficacy (e.g., De Castella and Byrne, 2015), perceived control, and attribution style (e.g., Schleider and Schroder, 2018). By the same token, implicit self-theories should not be confused with other similar-sounding theories, such as implicit knowledge of beliefs (Clements and Perner, 1994) implicit theories of the self (Greenwald et al., 2002) and implicit Theory of Mind (Clements, 2000).

Yeager and Dweck (2020), summarize implicit self-theory, as a theory that attempts to explain differences in individuals' responses to challenges or setbacks (p. 4). This concise, but perhaps far-too-simplistic summary notwithstanding, a large body of research indicates that these contrasting theories regarding the nature of human traits influence how people select and process information, form judgements, and act in consequence of these valuations (Dweck, 2013; Dweck and Sorich, 1999; Burnette et al., 2013; Plaks et al., 2001; Dweck and Yeager, 2019). In other words, implicit self-theories appear to compel people toward a theory-consistent experience of reality (Dweck, 2017; Priester and Petty, 2016).

For instance, implicit self-theories can predispose people toward divergent goals (Dweck and Leggett, 1988; Elliott and Dweck, 1988). The incremental theory orients people toward learning goals, which involve effort (e.g., mastering a challenging task), and the potential for skill acquisition (Dweck et al., 1995a; Dweck and Sorich, 1999). Additionally, incremental theorists tend to persist in the face of difficulty (Nussbaum and Dweck, 2008; Moser et al., 2011), view failure as a chance to learn (Dweck, 2009; Robins and Pals, 2002), and attribute low performance to low effort

[4] Implicit self-theories, however, may emerge through conscious experience (e.g., being praised in childhood for one's apparent fixed intelligence; Mueller and Dweck, 1998), that, over time, become entrenched in automatic, unconscious mechanisms that ultimately shape perception and behavior (for an overview see Dweck, 2017; Haimovitz and Dweck, 2017). As discussed later, these beliefs may also arise from momentary inducements (Chiu et al., 1997b).

(Jain et al., 2009; Molden and Dweck, 2006).

While an incremental theory is motivated toward learning goals, an entity theory is characterized by a fundamental propensity towards performance goals (Dweck and Bempechat, 1983; Elliott and Dweck, 1988), in which they seek to obtain approval and positive judgements from others (Dweck, 2013), through displays of supposedly inherent and fixed pre-existing competencies (Blackwell et al., 2007), and the avoidance of failure (Hong et al., 1997), even if this requires cheating (Blackwell et al., 2007) giving up, or withdrawing effort to hide potential inaptitude (Dweck and Leggett, 1988; Molden and Dweck, 2006). For the entity theorist, the exertion of effort is viewed as a sign of incompetence (Grant and Dweck, 2003), because if one has high competence, high effort is unnecessary (Murphy and Dweck, 2016b). Likewise, poor performance is considered to be a reflection of low ability or low intelligence (Blackwell et al., 2007; Hong et al., 1999).

Moreover, extant research has shown that while incremental theorists focus on the process of achievement, entity theorists focus on outcomes (Jain et al., 2009; Levy et al., 1998). An incremental theory has also been linked to increased positive emotions (e.g., happiness and excitement; see also Howell et al., 2016). According to some (e.g., Burnette et al., 2013), this is due, in part to the incremental theorists' proclivity to regard difficulties as learning opportunities (Dweck and Leggett, 1988; Nussbaum and Dweck, 2008; Mathur et al., 2014). Meanwhile, the entity theory has been associated with negative emotions (e.g., anxiety and anger; see also De Castella et al., 2013) partly a consequence of relating unfavorable outcomes with personal characteristics (Dweck et al., 1995a; Robins and Pals, 2002; Ommundsen et al., 2005; Miller et al., 2007).

As noted above, implicit self-theory orientation determines, to some degree, the information people select and pay attention to (Poon and Koehler, 2006). Specifically, individuals with an incremental theory focus on both belief-confirming and belief-disconfirming information (Plaks et al., 2001). Entity theorists on the other hand, mainly focus on information that confirms their beliefs (Plaks et al., 2005).

Importantly, an individual's implicit self-theory appears to exert a powerful influence on the attributions they make with respect to other people (Dweck et al., 1995a, 1993). For instance, entity theorists have been found to make snap judgements about the traits of others (Chiu et al., 1997a), even in the presence of insufficient information (Heslin et al., 2005). In keeping with their belief that traits are the direct cause of behavior (Plaks et al., 2001), entity theorists expect these personal qualities to be fixed, consistent in all situations, and indicative of future behavior (Dweck et al., 1993) even in the presence of contradictory information (Gervey et al., 1999). This view can extend beyond individual attributions to include group attributions as well (see Halperin et al., 2011; Rydell et al., 2007).

More generally, individuals who hold entity self-theories, in comparison to those who hold incremental self-theories, are prone to making stereotypical judgments (Levy et al., 1998; Plaks et al., 2001, 2005), and are inclined to higher degrees of attitude certainty (Petrocelli et al., 2010).

Conversely, incremental theorists make fewer trait attributions (Plaks et al., 2005; Levy and Dweck, 1998) and rather base their perceptions of people on situational and environmental factors (Dweck et al., 1993, 1995a) while expecting behaviors to vary across contexts (Levy et al., 1998).

Individuals with an incremental theory also exhibit increased preference for rehabilitation rather than punishment when deciding on how to address wrongdoing, than those with an entity theory, who by contrast favour punishment and retribution (Erdley and Dweck, 1993; Gervey et al., 1999), partly because they blame fixed (negative, in this case) traits for undesirable behavior (Chiu et al., 1997a; Tam et al., 2013; Quintanilla, 2011; Wurthmann, 2013). It is perhaps, no surprise, then that incremental theorists, more so than entity theorists, are inclined to exert effort in building interpersonal connections with others (Knee et al., 2003, 2004), and as such, have longer-lasting relationships (Knee, 1998).

It might be noted here that implicit self-theories appear to be domain-specific, in that, it is possible for an individual to simultaneously hold, say, an entity theory of intelligence and an incremental theory of personality (Hong et al., 1997; Hughes, 2015; Schroder et al., 2016). Furthermore, domain-specific self-theories correlate with same-domain outcomes. For example, implicit self-theories of personality are more relevant to interpersonal perceptions and behaviors than are implicit self-theories of intelligence (Schleider and Schroder, 2018).

Like any dominant paradigm, however, implicit self-theory has been subjected to some criticism (e.g., Burgoyne et al., 2020) in recent years. Primarily, with respect to issues of replication and generalizability (e.g., Glerum et al., 2019; Li and Bates, 2019). However, many of these criticisms have now been addressed (see Yeager and Dweck, 2020, reply).

Of interest as well is that these beliefs, although reasonably stable, are also readily labile (see Burnette et al., 2013, for review). Past research is consistent in showing that implicit self-theory can be primed and altered, at least temporarily (e.g., Miu and Yeager, 2015; Schleider and Weisz, 2016), using a variety of external stimuli such as articles (e.g., Chiu et al., 1997b), video media (e.g., Aronson et al., 2002) and advertising (e.g., Yorkston et al., 2010). For instance, Bergen (1991) induced participants to adopt either an incremental or entity theory by exposing them to a fictitious article that endorsed one or the other of these two orientations. In a similar manner, Mathur et al. (2013) successfully elicited different implicit self-theories by showing participants a list of synonyms for the words “fixedness” or “changeability”. Likewise, Jain and colleagues (2009) manipulated incremental beliefs by showing people television clips portraying characters that embody either an entity or an incremental view. According to Dweck et al. (1995a), this is made possible because most individuals’ hold both self-theories, albeit to varying degrees, therefore the less dominant theory may be accessed under priming conditions (Yorkston et al., 2010). As well, extant research has repeatedly demonstrated that experimentally primed self-theories’ bias preferences in a comparable manner to those found in studies measuring them as chronic orientations (see for example Molden and Dweck, 2006; Plaks et al., 2005; Chiu et al., 1997b; McConnell, 2001). Furthermore, the act of inducing implicit self-theories, allows researchers to assess their causal role, that is, to rule out the possibility that results may be explained by other variables related to measured levels of entity or incremental theory endorsement (see Yorkston et al., 2010; Park and John, 2012). Somewhat relatedly, meta-analyses of interventions targeting students’ implicit self-theories, have shown significant long-term improvements in academic performance, especially in students facing challenging transitions or those from low-income backgrounds (see Yeager and Walton, 2011; Paunesku et al., 2015; Yeager et al., 2019). Indeed, accumulating evidence suggests that such priming interventions can be reasonably simple, short, and inexpensive

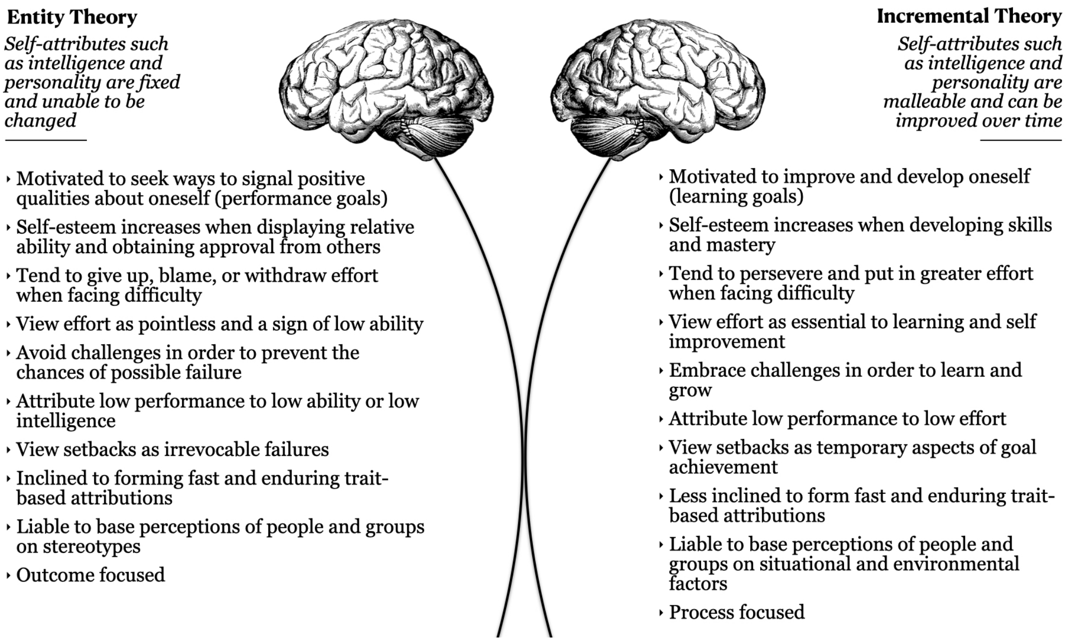


Figure 1.1: Diagram depicting the primary characteristics of entity and incremental theorists. *Note.* Differences in implicit self-theory are most striking in circumstances of challenge or difficulty, and when facing setbacks (own image).

(e.g., Paunesku, 2013; Bostwick and Becker-Blease, 2018).

In sum, a long tradition of research indicates that people's implicit self-theories about the adaptivity or stability of self-attributes such as intelligence and personality, have substantial effects on an array of human functioning (i.e., they can direct the perception and selection of information and drive expectations, motivations, and behaviors, among other things). This construal of human traits, although stable, can also be altered via exposure to selective information. Figure. 1.1 summarizes the primary characteristics of entity and incremental theorists.

Implicit Self-Theories Across Domains

Despite the overwhelming evidence linking implicit self-theories to a variety of perceptions, expectations, and behaviors, relatively few studies have investigated these beliefs outside the domain of academic and educational achievement (e.g., Ablard, 2002; Ablard and Mills, 1996; Blackwell et al., 2007), and apparently no empirical study has examined implicit self-theories in the context of HRI to any great extent (discussed later in this chapter). Nonetheless, converging evidence from this small but growing body of research indicates that implicit self-theories have downstream influences in a wide range of situations (Mathur et al., 2014) which suggests they may (potentially) apply to HRI issues.

In one set of studies, Yeager and colleagues (2013), for example, found that adolescents with an entity theory interpreted ambiguous provocations from (unknown) peers to be hostile in intent, which in turn lead them to express relatively violent desires for revenge (see also Yeager et al., 2011). This finding appears to support the notion that entity theorists often experience more negative affect than positive (e.g., King, 2016; Kim et al., 2018; King et al., 2012), and tend to make global trait judgements about others based on limited information (e.g., Gervey et al., 1999;

Heslin et al., 2005).

In recent research on implicit self-theories and financial decision making, individuals who held an incremental theory were marked by the tendency to prefer risk-seeking investments. By contrast, entity theorists appeared to prefer risk-averse investments (Rai and Lin, 2019). In addition, this research found that, at least in the context of financial decision making, incremental theorists are more prone to adopt a promotion-focus, whereas entity theorists may orient towards a prevention-focus (see also Montford et al., 2019). These findings are in line with prior work showing that individuals who adhere to incremental self-theories exhibit greater optimism, perseverance, and resilience in the face of setbacks (e.g., Mangels et al., 2006; Hong et al., 1999), and a future temporal focus (e.g., Price et al., 2017), in that they set goals to improve their abilities in the future (Dweck and Leggett, 1988). As contrasted with entity theorists, who have been repeatedly shown to be more sensitive to negative events (Robins and Pals, 2002; Elliott and Dweck, 1988) adopt poorer stress mitigation strategies (Doron et al., 2009), and hold a present temporal focus (Dweck and Leggett, 1988), that is they are more concerned with signalling their traits in the present (Sevincer et al., 2014).

Other studies have examined implicit self-theories regarding branding, advertising and consumer behavior (see Jain and Weiten, 2020, for an overview). For instance, Park and John (2010) found that entity theorists were particularly attracted to luxury brands (e.g., Victoria's Secret), in large part because such brands provide entity theorists with the means through which to signal positive self-traits (e.g., attractiveness). Contrastingly, incremental theorists were uninfluenced by such brands. Similarly, advertising appeals that highlight a brand's signalling abilities have been shown to be more popular with entity theorists, whereas incremental theorists systematically prefer advertising appeals that focus on aspects of self-improvement (see Park and John, 2012).

In the case of brand extensions, Yorkston and colleagues (2010) have found that individuals who hold an entity theory are less likely to accept a brand's attempt to extend into new and diverse product categories (e.g., Subaru extending into motorized scooters). The authors explain that these findings are a byproduct of entity theorists' belief that a brand's expertise, personality, and image is fixed and unable to change (see also Mathur et al., 2012).

Jain et al. (2009) demonstrated, similarly, that positively, rather than negatively framed advertising messages had a persuasive effect on incremental theorists, but had no influence on entity theorists, who, emphasizing outcome (Levy and Dweck, 1998; Mathur et al., 2013), perceived the alternatively framed messages as representative of the same result.

More recently, research in the realm of social media marketing has demonstrated that incremental theorists are more inclined to follow brands that they have never used before. In contrast, entity theorists are more likely to follow brands they already use (Song et al., 2019). This finding is in keeping with previous observations, showing that incremental theorists are motivated toward new possibilities, while entity theorists tend to confirm prior choices (e.g., Dweck, 2013) and seek to maintain the status quo (Quintanilla, 2011; Kam, 2011; Plaks and Stecher, 2007).

In the domain of trust-recovery, incremental theorists who had their trust violated in a repeated trust game, were quicker to regain trust following an apology and a promise to change, than were entity theorists (Haselhuhn et al., 2010). This finding falls in line with earlier research demonstrating that incremental theorists tend to believe in people's ability to change (e.g., Kammrath and Peetz, 2012), which in turn may motivate them to be more forgiving compared to entity theorists (Erdley and Dweck, 1993; Williams, 2015).

Contrarily, Haselhuhn and his colleagues (2008) have shown that in a negotiation game, entity theorists retained their trust in a negotiation partner even after they were made aware that their partner had deceived them. Incremental theorists, on the other hand, decreased their trust in their partner considerably once they learned of the deception. This finding mirrors the results of prior work showing that entity theorists are both quick to draw conclusions about others and resistant to updating their perceptions when faced with contrary information (e.g., Gervey et al., 1999; Heslin et al., 2005).

In the context of technology acceptance, Sharifi and Palmeira (2017), demonstrated that incremental theorists showed more favorable responses to a technology product (a smart ring) perceived as complex. Whereas entity theorists indicated aversion toward the product. The same product, however, when perceived as simple, induced similar reactions from both incremental and entity theorists. These results in particular, seem consistent with other research showing that incremental theorists gravitate towards effortful engagement (Dupeyrat and Mariné, 2005) in order to learn new tasks, which may predispose these individuals to accept technology (Solberg et al., 2020) more broadly, and radical technological innovations (Hafeez, 2019), more specifically.

In one study on information technology continuance, Fong et al. (2018) reported that incremental theorists were more inclined than entity theorists to use, and continue to use, an app to make hotel reservations. This finding appears consistent with prior research showing entity theorists tend to hold negative effort beliefs (e.g., Murphy and Dweck, 2016b; Hong et al., 1999; Knee, 1998), and are lower in personal innovativeness (Aldahdouh et al., 2018), which is a determinant of technology adoption in general (Jin, 2013).

In another study, conducted by Kim et al. (2018), entity theorists showed greater aversion towards an anthropomorphized helper than a non-anthropomorphized helper in an online learning task. Incremental theorists, by contrast, were not affected by anthropomorphism. This finding appears to indicate that anthropomorphic features may precipitate entity beliefs about receiving help from a person, rather than a software program, which in turn, can signal incompetence on the part of the entity theorist (Dweck and Leggett, 1988; Dweck, 1986; Elliott and Dweck, 1988).

As already noted, two studies have (partially) explored implicit self-theories with respect to child–robot interaction. In particular, Park et al. (2017) examined the effect of a child-friendly robot that expressed a growth mindset whilst interacting with young children during a puzzle-solving game. They showed that children were likely to adopt the robot's growth mindset as expressed in their self-reported beliefs and task-performance measures following a challenge.

Similarly, Davidson and colleagues (2021) conducted an unsupervised longitudinal study whereby children interacted with either a Computer Aided Learning (CAL) system that administered effort-focused praise via headphones or one that delivered the same praise via a social robot. This study found that children who received effort-related praise from the social robot had a significant increase in their incremental theory score, compared to those who received praise from the CAL system.

Together, these two studies shed at least some light on the role of implicit self-theory in HRI. Nevertheless, they are limited by the fact that the researchers examined only a single implicit self-theory dimension (incremental). In addition, these studies did not investigate the role of implicit self-theories in shaping attributions and behavior per se, therefore people's evaluations of a robot post-task were not examined. On top of that, these studies had relatively small samples (44 participants or fewer), which were composed of children, and the latter reported a few inclusive or rather, weak findings. Thus, it is currently unknown how these beliefs relate to HRI.

Considering the impact of implicit self-theories in diverse domains such as intelligence, personality, relationships, stereotypes, negotiation, and trust recovery, as well as its influence in the realms of consumer behavior, technology acceptance, and financial decision making, it seems reasonable to presume that implicit self-theories might be a central factor in explaining phenomena of importance to the sphere of HRI research. However, it is important to note that this is by no means a foregone conclusion. Indeed, social robots, owing in part to their humanlike embodiment (see Ferrari et al., 2016; Haring et al., 2018), perceived agency and experience (see Gray and Wegner, 2012), social capabilities (see Collins, 2019), and capacity for eliciting affective responses (see Damiano and Dumouchel, 2018), have been identified as a distinctly different class of product (see de Graaf et al., 2016; Severson and Carlson, 2010), and one that defies clear categorization (see Kahn et al., 2011; Strait et al., 2019). This has led several researchers (e.g., de Graaf et al., 2019; Damholdt et al., 2020) to argue that it is erroneous to accept, without question, that the fundamental propositions of the social sciences will invariably apply to social robotics and the field of HRI (see also Wullenkord and Eyssel, 2020). In this view, empirical research is needed to substantiate whether implicit self-theories influence people's perceptions and responses to social robots.

1.3 Objectives

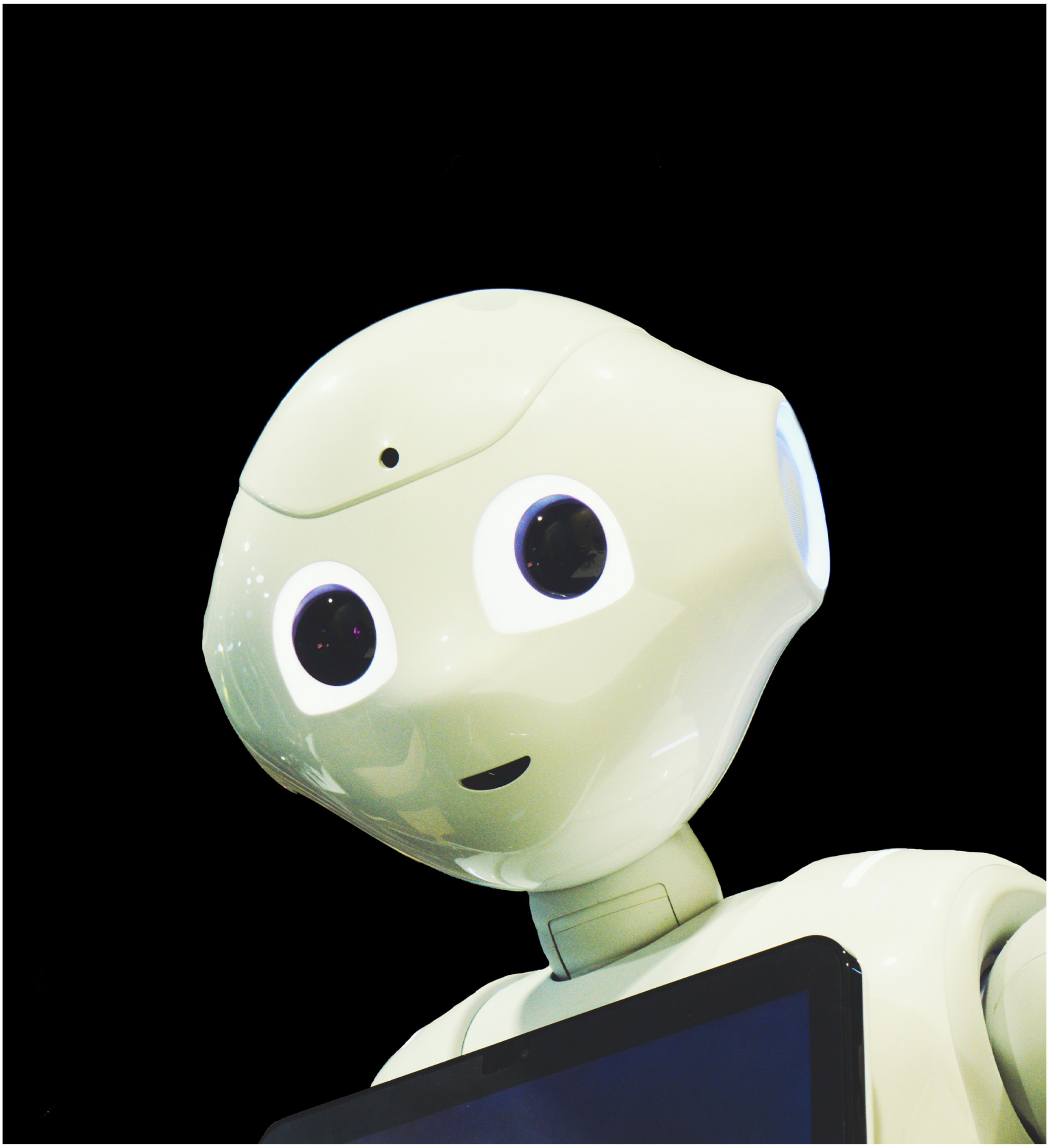
The objective of this research is to examine the proposition that implicit self-theories represent an important, and perhaps, even meaningful variable that influences the manner in which people perceive, evaluate, and respond to, social robots. To test this proposition, three studies investigated the role of implicit self-theories in relation to several issues relevant to HRI.

In the following chapters: Study 1 (Chapter 2) seeks to establish whether implicit self-theories influence individuals' perceptions of emerging social robots developed for everyday use. Then, Study 2 (Chapter 3) tests whether implicit self-theories and robot-delivered praise interactively effect people's evaluations of a social robot after success on a challenging task. Finally, Study 3 (Chapter 4) aims to complement Study 2 by testing the relationship between implicit self-theories and people's responses to an apparently intelligent autonomous robot after perceived failure.

Across studies, different scenario media are used, in conjunction with different robot types. Not to mention, different assessments of implicit self-theories are investigated on a variety of dependent variables. On top of this, experimental scenarios conducive to implicit self-theories (e.g., success/failure, challenges/setbacks) are tested.

Chapter II

Study 1



II

Study 1: The Influence of Implicit Self-Theories on People's Perceptions of Social Robots Developed for Everyday Use

Material covered in this Chapter has previously been published in Allan, D. D., Vonasch, A. J., and Bartneck, C. (2022b). The doors of social robot perception: The influence of implicit self-theories. *International Journal of Social Robotics*, 14(1):127–140

2.1 Introduction

This chapter describes the results of Study 1, which aimed to provide initial evidence that implicit self-theories exert influence on people's perceptions of emerging[1] social robots developed for everyday use. The structure of this chapter is organized as follows: Section 2.2 provides a brief overview of the study and presents the hypotheses. This is followed by Section 2.3 which describes the method used. After this, Section 2.4 summarizes the results, and Section 2.5 identifies some limitations and opportunities for future research. The chapter ends with a brief conclusion in Section 2.6.

2.2 Overview and Hypotheses

As detailed in the previous chapter, there is a dearth of research regarding the application of implicit self-theories in HRI. Thus, preliminary evidence is needed to substantiate whether implicit self-theories influence people's perceptions, eval-

[1]Social robots can be considered emerging technologies in that they are technologies either still in development (Conger et al., 2013), under-exploited, not mature (Einsiedel, 2009), or that most people have little to no experience with (Naneva et al., 2020).

uations, and responses to social robots.[2] Accordingly, Study 1 aimed to provide this critical data. Recall from Chapter 1 that, entity theorists, but not incremental theorists, tend to show more risk aversion (Rai and Lin, 2019), a preference for maintaining the status quo (Quintanilla, 2011; Kam, 2011; Plaks and Stecher, 2007), and less favorable reactions towards novel technology products (Sharifi and Palmeira, 2017; Fong et al., 2018). Furthermore, entity theorists, more so than incremental theorists, are low in personal innovativeness (Aldahdouh et al., 2018), which in turn, is a negative predictor of social robot acceptance (De Graaf et al., 2015). Therefore, the following hypotheses were posited:

Hypothesis 1 (H1): Entity (vs. incremental) theorists will exhibit more (vs. less) robot anxiety.

Hypothesis 2 (H2): Incremental (vs. entity) theorists will exhibit more (vs. less) attitudinal acceptance toward social robots.

Hypothesis 3 (H3): Entity (vs. incremental) theorists will exhibit less (vs. more) intentional acceptance[3] toward social robots.

Related to robot acceptance, a pivotal question in HRI is which role do people expect a social robot to perform?

Past research has shown that people consider the role of a social robot to be that of an assistant, domestic tool, or servant (Dautenhahn et al., 2005; Ray et al., 2008; Takayama et al., 2008; Ezer, 2008; Cagiltay et al., 2020), rather than as a social companion (Bernotat and Eyssel, 2018; de Graaf et al., 2019). What's more, there is evidence suggesting that individual difference factors may play an important role in determining the different functions that people expect a social robot to fulfill (see Copleston and Bugmann, 2008; Ezer, 2008; Nomura et al., 2009; Eyssel and Hegel, 2012; Bernotat and Eyssel, 2018). Interestingly, a recent paper on implicit self-theories (Han et al., 2019) reported that participants with different self-theories responded disparately to advertising messages for a travel brand anthropomorphized as a partner (vs. servant). Specifically, entity theorists responded more favorably to a brand anthropomorphized as a servant, whereas incremental theorists preferred a partner brand.

It seems plausible that entity theorists' desire for effortlessly attained benefits (Dweck and Leggett, 1988; Elliott and Dweck, 1988) and partiality for signaling their superiority (Dweck, 2017), led them to be more accepting of a brand anthropomorphized as a servant. Likewise, incremental theorists' preference for process and effort (Hong et al., 1997; Levy and Dweck, 1998), most probably resulted in

[2]As discussed in Section 1.1, a growing literature suggests that a person's views and evaluations of social robots predict their acceptance and intention to use them, and that this relationship is shaped by underlying psycho-social variables. Correspondingly, individuals' perceptions and responses toward social robots are measured on facets/dimensions of robot acceptance (e.g., robot anxiety, attitudinal acceptance, intentional acceptance; Krägeloh et al., 2019; Zimmerman et al., 2022).

[3]Following (Bernotat and Eyssel, 2018), attitudinal and intentional robot acceptance were assessed as different dimensions of robot acceptance. According to Ezer (2008), attitudinal robot acceptance refers to one's positive beliefs regarding a robot more generally, whereas intentional acceptance refers to an individual's intention to purchase or use a robot.

more favorable responses toward the brand-as-partner option. Based on this rationale, then, it would be of interest to explore whether this relationship exists in regard to social robots.

Hence, an additional aim of Study 1 was to extend beyond this work, and build on previous HRI findings (i.e., preferred robot role classifications), to examine the relation between the implicit self-theory held by participants and their evaluations of a social robot presented as either a collaborative assistant[4] or a personal servant. Formally, it was hypothesized that:

Hypothesis 4 (H4): Incremental (vs. entity) theorists will evaluate a robot described as a collaborative assistant (vs. servant) more (vs. less) favorably.

One must bear in mind, however that robot role preference was examined, specifically, in relation to intentional robot acceptance (e.g., Bernotat and Eyssel, 2018). Most notably, this preference is not particularly relevant to the subsequent set of studies, and therefore will not be examined or explored further.

It might be appropriate to interject here, that as implicit self-theories exist on a continuum (described in Section 1.2), they are, in turn, measured and analysed on a continuous scale[5] as is was the case in Study 1.

Consistent with the recommendation of Wullenkord and Eyssel (2020), who argue that HRI researchers should consider covariate influences, when initially employing knowledge from the social sciences, measures of gender (Eyssel et al., 2012), age (Kuo et al., 2009), education (Eurobarometer, 2012), media exposure to science fiction (Sandoval et al., 2014), technology commitment (Halperin et al., 2011), and prior robot experience (Bartneck et al., 2007) were assessed. These variables have been routinely shown to impact individuals' perceptions of robots (see Schermerhorn et al., 2008; Enz et al., 2011; Bartneck et al., 2020). Therefore, they were considered as covariates.

2.3 Method

The foregoing hypotheses were tested in an online human-subject experiment, which was compiled and hosted on Qualtrics. As indicated in Section 1.3 all hypotheses, design, planned sample sizes, and the analysis plan for Study 1 were preregistered and can be accessed via the Open Science Framework (OSF): https://osf.io/ymnbe/. There were several deviations from the preregistration. First, age was not included as a covariate. This deviation occurred because age was not recorded in the dataset due to a technical error. Second, the preregistered analyses were not conducted on or with the predicted interaction (implicit self-theory + robot role) and therewith Hypothesis 4 was not directly assessed. The reason for this is that the manipulation check failed to show that the manipulation of robot role was successful (described in

[4]For the purposes of Study 1, the role of "assistant" was used rather than "partner." The rationale for this is the empirical evidence on robot role, which has repeatedly shown that people distinguish between the two roles (servant vs. assistant). In contrast, there appears to be much less evidence for partner role type perceptions in the HRI literature (e.g., Dautenhahn et al., 2005)

[5]Robins and Pals (2002) have noted, that while implicit self-theories are analysed as a continuous variable,"to facilitate communication of the findings we often describe the results in terms of differences between Entity and Incremental theorists, reflecting individuals who are relatively high vs. low on the Entity scale" (p. 319).

Section 2.4.1), therefore, the predicted interaction could not be established. Thus, these analyses were no longer appropriate, much less meaningful. Part of this protocol was however, performed as a post-hoc exploratory analysis, the results of which are publicly available at `https://osf.io/abkwp/`. Successively, other exploratory analyses were conducted which were not specified in the preregistration. The research protocol was approved by the Human Research Ethics Committee of the University of Canterbury (HEC 2019/53/LR-PS).

2.3.1 Recruitment

Participants were recruited through Amazon's Mechanical Turk (MTurk) crowdsourcing platform. On Mturk "Requesters" or employers provide details of a particular task, also known as "Human Intelligence Task (HIT)" that they want completed. Those interested, known as "Workers" can opt to complete such HITs. Subsequently, Workers are compensated so long as the completed HIT is considered to be of a satisfactory quality as defined by the Requester. The use of MTurk for the purpose of running online studies, is an established and widely used practice in HRI research (Hoffman and Zhao, 2020; Feil-Seifer et al., 2020), owing in large part to the relatively fast, reliable and inexpensive data it produces (Bartneck et al., 2015; Buhrmester et al., 2016; Huff and Tingley, 2015; Kees et al., 2017). Participation in Study 1 was restricted to those who resided in the United States, had completed >50 surveys with prior HIT approval ratings of 98% or greater (Robinson et al., 2019). An a priori power analysis showed that the sample size (n=251) was necessary to detect a medium effect at a power of 0.95 and an alpha level of .05.[6]

2.3.2 Design

Study 1 used an implicit self-theory (continuous)×2 (role: Servant, assistant) between-subjects design.

2.3.3 Participants

A total of 251 participants agreed to take part in the study for a reimbursement of $1.50 in MTurk credit. Four participants who either failed to pass the attention check (Robinson et al., 2019), the validity indicator (Chmielewski and Kucker, 2020), or who asked to withdraw their data, were excluded from the dataset before any analyses were conducted. Out of 247 participants in the final sample, 45.34% were female. Most participants (65.47%) had undergraduate education (some college education) while 15.38% had postgraduate degrees and 9.72% had high school education. All participants provided consent before beginning the study.

[6]No comparable study could be located on which to base the expected effect size, therefore the widely used sampling perspective suggested by (Schönbrodt and Perugini, 2013), was adopted. In this approach, a minimum of 250 participants is considered preferable in order to obtain stable (i.e., significant but also accurate) correlations (for thorough explanations see Schönbrodt and Perugini, 2013; Kelley and Maxwell, 2003; Maxwell et al., 2008). That said, however, at pre-registration the required sample size was incorrectly entered (based on an earlier ANCOVA test). The correct required sample is n=95 to detect a medium effect for regression analyses at a power of 0.95 and an alpha level of .05. Fortunately, the misentered required sample size was greater than the correct required sample size, resulting in the study having sufficient power for a more precise and stable estimate as initially intended.

2.3.4 Procedure

After reading the instructions, participants were asked to complete a measure of implicit self-theory (Levy et al., 1998), described below in Section 2.3.5, after which they were randomly assigned to read one of two written descriptions briefly describing a new generation of social robots modified from (de Graaf et al., 2019).
Those in the assistant condition read:

> Social robots are created in such a way that they can operate independently as collaborative assistants in everyday environments, such as your home. Social robots can understand everyday social situations and react according to human social norms. Regarding social situations includes conversations between people as well as how we ought to behave in the presence of other people. Moreover, a social robot will work with you, helping you to achieve your goals, ultimately performing the role of a collaborative assistant. Social robots are able to communicate with us in a humanlike way through speech interactions with supportive gestures and facial expressions. Because a social robot is a collaborative assistant it can assist you as you go about accomplishing your daily tasks.

By contrast, participants in the servant condition read that:

> Social robots are created in such a way that they can operate independently as personal servants in everyday environments, such as your home. Social robots can understand everyday social situations and react according to human social norms. Regarding social situations includes conversations between people as well as how we ought to behave in the presence of other people. Moreover, a social robot works for you, executing tasks on your behalf, and ultimately performs the role of a personal servant. Social robots are able to communicate with us in a humanlike way through speech interactions with supportive gestures and facial expressions. Because a social robot is a servant it can do chores in and around the home based on your preferences.

A written description rather than video or pictures was used, as previous research suggests that a robot's appearance can affect individuals' perceptions regarding its function (Goetz et al., 2003; Walters et al., 2009) which may, in turn, result in unfavorable judgments (Haring et al., 2018; Reich-Stiebert et al., 2019). Figure. 2.1 depicts a wide array of commercially available social robots of varying form (to be clear, participants were not shown this image it is included here for illustrative purposes). Resultantly, written descriptions are universally accepted and used in HRI studies (Reich-Stiebert and Eyssel, 2015; de Graaf et al., 2019; Stapels and Eyssel, 2021). Upon reading the manipulation text, participants completed the dependent measures and provided demographic information. Finally, participants were debriefed and compensated $1.50 for their participation.

2.3.5 Measures

Implicit Self-Theory

Implicit self-theory was assessed using an 8-item measure from (Levy et al., 1998). On a 6-point scale (1 = strongly disagree, 6 = strongly agree), participants indicated

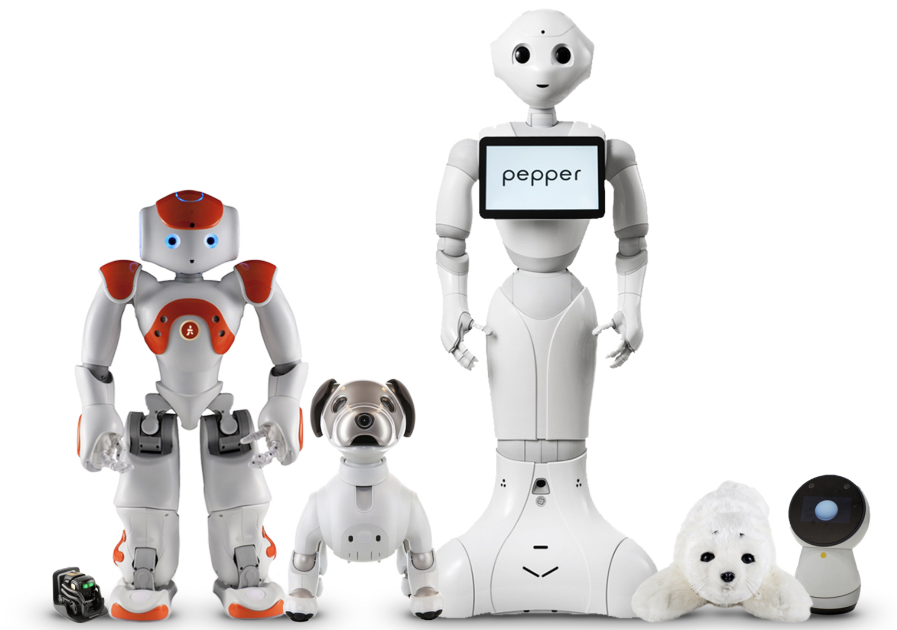


Figure 2.1: A selection of Social Robots varying in aesthetic form. From left to right: Vector, Nao, Aibo, Pepper, Paro, and Jibo. Images taken from `https://robots.ieee.org/robots/`

their agreement with four items representing entity beliefs (e.g., "Everyone is a certain kind of person, and there is not much that they can do to really change that"), and four items representing incremental beliefs (e.g., "People can change even their most basic qualities"). The four incremental items were reverse-scored and averaged with the four entity items to create a single measure of implicit self-theory with higher scores (4.0 and above) reflecting an incremental self-theory, and lower scores (below 3.0) corresponding to more of an entity self-theory. Respondents with scores between 3.0 and 4.0 are classified as having an undefined self-theory (Levy et al., 1998). It is crucial to note that this self-report scale is the most widely used and well validated method of measuring domain-general implicit self-theories (e.g., Levy and Dweck, 1998) in experimental settings (Mathur et al., 2016). As such, this measure has demonstrated efficacy in a range of domains including those related to stereotypes (e.g., Levy et al., 1998), advertising (e.g., Mathur et al., 2013), financial decision-making (e.g., Montford et al., 2019), performance appraisals (e.g., Heslin et al., 2005) and consumer behavior (e.g., Yorkston et al., 2010). The Cronbach's alpha was 0.95.

Manipulation Check

Two items modified from (Kim and Kramer, 2015) assessed the extent to which participants perceived a social robot as a servant (e.g., "A Social Robot is like a servant to the consumer") versus as an assistant (e.g., "A Social Robot is like an assistant to the consumer"). Agreement with items was indicated using a 7-point Likert scale (1=strongly disagree, 7=strongly agree).

Robot Anxiety (H1)

Robot anxiety was assessed with the 14-item NARS (Negative Attitudes toward Robots Scale; Nomura et al., 2004). Respondents indicated agreement with items (e.g., "I would hate the idea that robots or artificial intelligences were making judgments about things") on a 5-point scale (1=strongly disagree, 5=strongly agree) with higher scores indicating more negative attitudes. This scale has been used in previous research to measure robot anxiety (e.g., Bernotat and Eyssel, 2018; Bartneck et al., 2007; Xia and LeTendre, 2020). The Cronbach's alpha was 0.89.

Attitudinal Robot Acceptance (H2)

Five items adapted from (Ninomiya et al., 2015) assessed participants' likelihood of acceptance. Agreement with items (e.g., "It is good if a social robot can do the work of a human," and "I would want to boast that I have a social robot in my home"), was indicated on a 7-point scale (1=strongly agree, 7=strongly disagree). The Cronbach's alpha was 0.88.

Intentional Robot Acceptance (H3)

Four items from (de Graaf et al., 2019) measured participants' intentions to use a robot (e.g., "Assuming I have a robot, I will frequently use it in the future," and "I think a social robot would be useful to me"). As well, one item asked participants to indicate their likelihood of purchase (i.e., "Assuming a social robot is affordable, I will likely purchase one in the future"). The five items were measured on a 7-point scale (1 = strongly agree, 7 = strongly disagree). The Cronbach's alpha was 0.95.

Robot Evaluation (H4)

A three-item, seven-point scale modified from (Han et al., 2019) assessed participants' evaluations of the social robot (i.e.,"very unfavorable/very favorable," "very bad/very good," and "very negative/very positive"). Higher scores correspond to higher endorsement. The Cronbach's alpha was 0.94.

Prior Robot Experience

Three items adapted from (Reich-Stiebert and Eyssel, 2015) measured participants' prior experience with robots (e.g., "Have you ever used, or are you currently using a robot at home or at work?"). A prior robot experience score was computed by taking the mean response to the three items. The Cronbach's alpha was 0.65.[7]

Media Exposure to Science Fiction

One item adapted from (Liang and Lee, 2017) evaluated participants' media exposure to science fiction. Respondents indicated agreement with this item (i.e., "How often do you watch television shows and movies related to science fiction and fantasy?") on a 7-point scale (1=very often, 7=never).

[7] This questionable internal consistency is not particularly surprising, given the low number of items available. As Tavakol and Dennick (2011) opine, "[A] high coefficient alpha does not always mean a high degree of internal consistency. This is because alpha is also affected by the length of the test. If the test length is too short, the value of alpha is reduced" (p. 53).

Table 1.1: Mean scores on dependent measures (N = 246).

Measure	Mean	*SD*	α
Robot Anxiety	2.68	0.75	0.89
Attitudinal Robot Acceptance	4.15	1.44	0.88
Intentional Robot Acceptance	5.01	1.50	0.95
Robot Evaluation	5.56	1.29	0.94
Prior Robot Experience	0.95	1.00	0.65
Media Exposure to Science Fiction	4.54	1.58	
Technology Commitment	3.79	0.60	0.81

SD represents standard deviation; α represents Cronbach's Alpha reliability statistic value.

Technology Commitment

Technology Commitment was assessed with the English version of the Technology Commitment scale (Neyer et al., 2012). The scale consisted of 12 items, assessing interest in, and acceptance of new technology (e.g., "I am very curious about new technical developments"). Agreement with items was indicated on a 5-point scale (1=strongly agree, 5=strongly disagree). The Cronbach's alpha was 0.81.

2.4 Results

2.4.1 Data Analysis

Four basic regression (multiple linear regression) analyses were performed with implicit self-theories (entity vs. incremental) as the predictor variable[8]. Robot anxiety, attitudinal robot acceptance, intentional robot acceptance, and robot evaluation, were used as dependent variables. Table 1.1 displays the mean scores on the dependent measures. The regression model was run separately for each hypothesis.

In every regression, individual differences in prior robot experience, media exposure to science fiction, technology commitment, as well as gender and education were included as covariates. Subsequently, the same series of regression analyses were performed without covariates. As well, separate regressions and simple slope analyses were performed as part of an exploratory analysis.

The predictor variable was centered prior to being entered into the model (Li et al., 1998). Regression diagnostic tests were then conducted for examining the validity of the model. The initial diagnostic analysis identified one extreme outlier due to high leverage in Cook's distance. With this outlier removed, the statistical assumptions of the model were met (Aiken et al., 1991).

[8]As described in Section 2.3.5, higher (vs. lower) scores on the implicit self-theories measure indicate more (vs. less) of an incremental (vs. entity) theory. Therefore, positive (vs. negative) statistically significant effects indicate an association between incremental (vs. entity) theory and the dependent variable of interest.

Manipulation Check

Independent sample t-tests results indicated that participants in the servant condition scored significantly ($t(244) = 6.09,\ p < .001\ d = 0.77$, (95% CI [0.50, -1.04]) higher on servant role perceptions ($M = 5.57,\ SD = 1.36$) than those in the assistant condition ($M = 4.45,\ SD = 1.53$). However, participants in the assistant condition did not score significantly ($t(244) = -1.50,\ p = .134,\ d = -0.19$, (95% CI [-0.44, -0.05]) higher on assistant role perceptions ($M = 5.81,\ SD = 0.917$) than those in the servant condition ($M = 5.59,\ SD = 1.34$). Therefore, the manipulation of robot role type was not entirely successful. Hence, and as discussed in Section 2.3 the planned test of Hypothesis 4 could not be performed, as this required a successful manipulation of robot role.

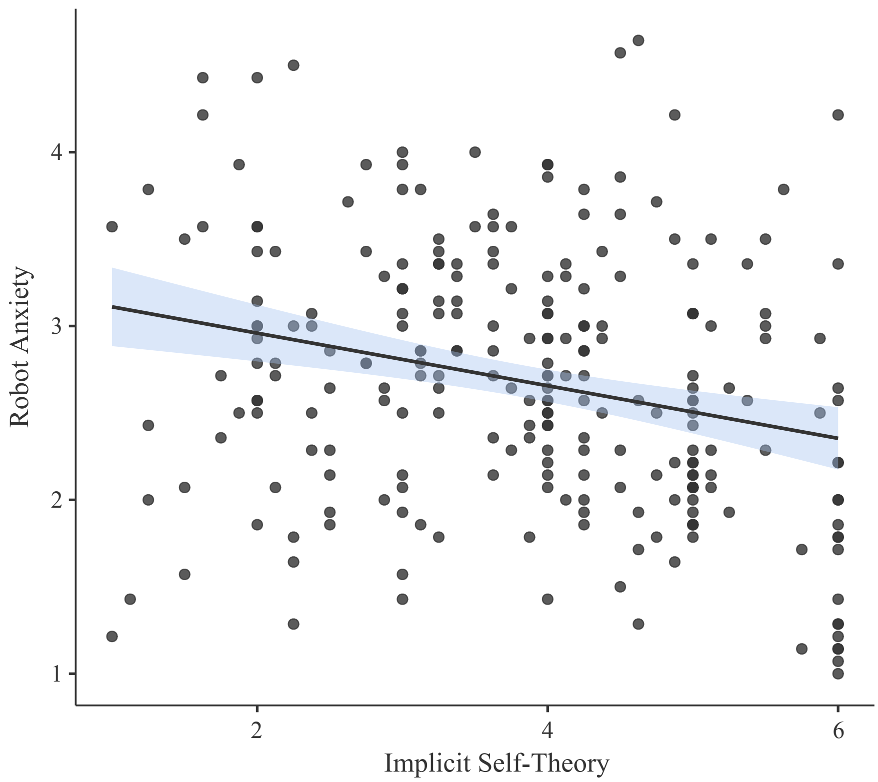


Figure 2.2: Scatterplot with the correlation line between implicit self-theory (lower scores indicate more of an entity theory) and robot anxiety. Confidence bands indicate 95% confidence intervals (CIs).

Robot Anxiety (H1)

According to H1, entity theorists were expected to report more robot anxiety than incremental theorists (H1). A statistically significant main effect of implicit self-theory on participants' ratings of robot anxiety ($B = -0.08,\ SE = 0.03,\ t = -2.48,\ p = .014,\ \beta = -0.13$, (95% CI [-0.24, -0.02]) was observed. That is, having an entity (vs. incremental) self-theory was significantly correlated with more (vs. less) robot anxiety. Consequently, H1 was supported. Regarding the covariates, a statistically significant main effect of participants' media exposure to science fiction, ($B = -0.06,\ SE = 0.02,\ t = -2.35,\ p = .019,\ \beta = -0.13$, (95% CI [-0.24, -0.02]), and a significant main effect of technology commitment ($B = -0.57$, $SE = 0.07$, $t = -7.96$, $p < .001$ $\beta = -0.46$ (95% CI [-0.46, -0.57]) was observed. Additionally, participants' education level ($B = -0.33$, $SE = 0.16$, $t = -2.00$, $p = .046$, $\beta = -0.43$, (95% CI [-0.85, -0.00]) had a significant effect on their level of robot anxiety. The largest effect size in the model was attributed to technology commitment ($\beta = -0.46$). No other significant effects were revealed.

Attitudinal Robot Acceptance (H2)

According to H2, it was expected that incremental theorists would indicate more attitudinal robot acceptance than entity theorists. A statistically nonsignificant main effect of implicit self-theory ($B = -0.07$, $SE = 0.07$, $t = 1.03$, $p = .303$, $\beta = -0.06$, (95% CI [-0.06, -0.19]) was observed. That is, having an incremental (vs. entity) self-theory was marginally correlated with more (vs. less) attitudinal robot acceptance. Thus, H2 cannot be said to be entirely supported. Participants' media exposure to science fiction tended to influence their attitudinal robot acceptance ($B = 0.14$, $SE = 0.06$, $t = 2.39$, $p = .017$, $\beta = -0.15$, (95% CI [0.02, 0.02]). Moreover, participants' self-rated technology commitment had a statistically significant effect on their attitudinal robot acceptance ($B = 0.49$, $SE = 0.16$, $t = 3.04$, $p = .003$, $\beta = -0.20$, (95% CI [0.07, 0.33]). The largest effect size in the model was observed for technology commitment ($\beta = 0.20$). No other significant effects were revealed.

Intentional Robot Acceptance (H3)

According to H3, entity theorists were expected to indicate less intentional robot acceptance than incremental theorists. A statistically nonsignificant main effect of implicit self-theory, ($B = 0.06$, $SE = 0.07$, $t = -0.86$, $p = .386$, $\beta = -0.20$, (95% CI [0.07, 0.33]) was observed. That is, having an entity (vs. incremental) self-theory was marginally correlated with less (vs. more) intentional robot acceptance. Thus, H3 was not completely supported. In regard to the covariates, participants' intentional robot acceptance was influenced statistically significantly by their media exposure to science fiction ($B = 0.13$, $SE = 0.06$, $t = 2.23$, $p = .026$, $\beta = -0.14$, (95% CI [0.01, 0.27]). In addition, participants' self-rated technology commitment had a significant effect on their intentional robot acceptance ($B = 0.68$, $SE = 0.16$, $t = 4.17$, $p < .001$, $\beta = -0.27$, (95% CI [0.14, 0.40]). The largest effect size in the model was seen for technology commitment ($\beta = 0.27$). No other significant effects were revealed.

Robot Evaluation (H4)

According to H4, entity theorists were expected to prefer a robot described as a servant (vs. assistant), whereas incremental theorists were expected to prefer a robot described as an assistant (vs. servant). However, as explained above, this interaction was not tested directly, and therefore by extension, neither was Hypothesis 4, due to the manipulation check failure. However, a statistically nonsignificant main effect of implicit self-theory, on participants' robot evaluation ($B = -0.07$, $SE = 0.06$, $t = -1.22$, $p = .223$, $\beta = -0.07$, (95% CI [-0.04, -0.04]) was observed. That is, having an incremental (vs. entity) self-theory was marginally correlated with a less (vs. more) favorable robot evaluation. Considering the covariates, a statistically significant main effect of participants' technology commitment ($B = 0.48$, $SE = 0.14$, $t = -3.40$, $p < .001$, $\beta = -0.22$, (95% CI [0.09, 0.35]) and gender, ($B = -0.33$, $SE = 0.16$, $t = -2.03$, $p = .043$, $\beta = -0.26$, (95% CI [-0.51, -0.51]) was observed. Gender had the largest effect size in the model at ($\beta = -0.26$). No other significant effects were revealed.

Exploratory Analysis

To gain a further insight into the relationship between participants' implicit self-theory and their perceptions of emerging social robots, the same regression analyses was run without covariates. Unsurprisingly, results show that implicit self-theory remains significant for robot anxiety ($B = -0.15$, $SE = 0.03$, $t = 4.14$, $p < .001$, $\beta = -0.25$, (95% CI [-0.37, -0.134]). However, a statistically significant main effect of implicit self-theory on robot evaluation ($B = -0.13$, $SE = 0.06$, $t = -2.09$, $p = .038$, $\beta = -0.13$, (95% CI [0.00, 0.357]) appears. That is, having an entity (vs. incremental) self-theory was significantly correlated with a less (vs. more) favorable robot evaluation. Moreover, attitudinal robot acceptance ($B = -0.13$, $SE = 0.07$, $t = -1.86$, $p = .064$, $\beta = -0.11$, (95% CI [-0.00, 0.244]), and intentional robot acceptance ($B = -0.14$, $SE = 0.07$, $t = -1.89$, $p = .061$, $\beta = -0.12$, (95% CI [-0.00, 0.245]), narrowly failed statistical significance. The largest effect size was found with robot anxiety ($\beta = -02$) and the smallest effect size was found with attitudinal robot acceptance ($\beta = 0.11$). No other significant effects were revealed.

A secondary exploratory analysis, was conducted to examine more closely the links between implicit self-theory and robot role. To do so, a similar model was performed with the manipulation check scores (servant vs. assistant), with implicit self-theory (entity vs. incremental), and their interaction as independent variables. The same dependent variables were used as above (i.e., robot anxiety, attitudinal robot acceptance, intentional robot acceptance, and robot evaluation). The results revealed a significant interaction ($B = -0.05$, $SE = 0.02$, $t = -2.85$, $p = .005$) of implicit-self-theory and robot-as-servant role with robot evaluation. To probe the nature of this interaction, a simple slope analysis was performed for "high" (1 SD above the mean), "medium" (mean), and "low levels" (1 SD below the mean) of robot-as-servant role perceptions ($M = 5.80$, $SE = 0.15$, 95% CI = [5.49, 6.11]), were not associated with more favorable evaluations than low levels of robot-as-servant role perceptions ($M = 5.69$, $SE = 0.15$, 95% CI = [5.40, 5.98]). However, for the entity theorists, higher levels of robot-as-servant role perceptions ($M = 5.74$, $SE = 0.15$, 95% CI = [5.43, 6.04]) were associated with higher robot evaluations than low levels of robot-as-servant role perceptions ($M = 4.96$, $SE = 0.17$, 95% CI = [4.62, 5.30]). Thus, entity theorists responded more favorably to a robot positioned as a servant. This finding, though not definitive, provides some partial support for H4. In contrast to this, no significant interaction effects were found for the other dependent variables.

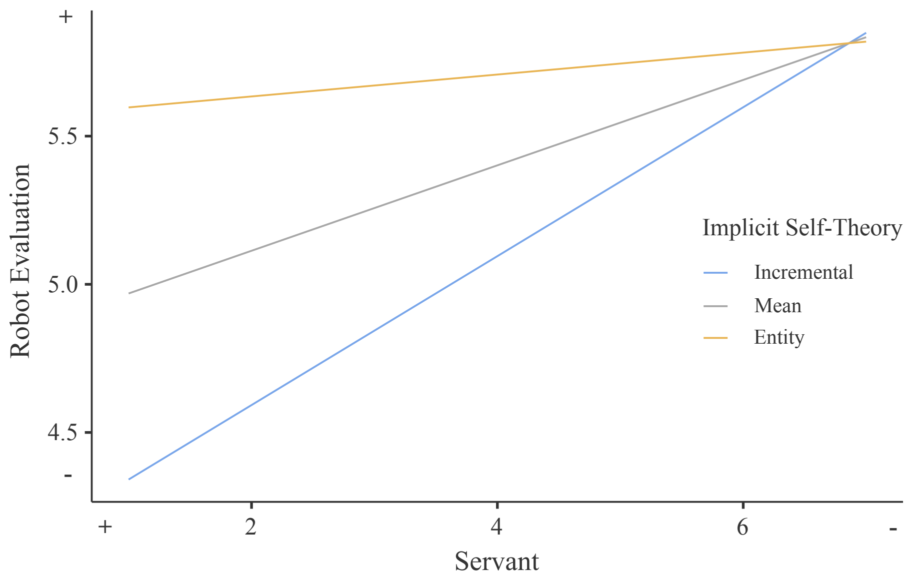


Figure 2.3: Linear regression lines depicting a significant interaction between implicit self-theory and robot-as-servant role in predicting robot evaluation.

2.5 Discussion

Study 1 provides initial evidence that one's implicit self-theory orientation influences one's perceptions of emerging social robots developed for everyday use. In particular, entity theorists showed greater robot anxiety than incremental theorists. Notably, this was true even when controlling for a number of important covariates.

Importantly, although the regression results for Hypothesis 2 and Hypothesis 3, followed the expected pattern, both narrowly failed statistical significance. Consequently, these hypotheses were not supported.

The above notwithstanding, an exploratory analysis showed that the effect of implicit self-theory on robot evaluation was significant when not controlling for covariates. That is incremental theorists, rated both robot types (i.e., servant and assistant) more favorably than entity theorists. This finding falls in line with recent research, showing that the incremental theory is positively associated with technology acceptance (Fong et al., 2018) and use (Solberg et al., 2020) more generally, and, more precisely, the endorsement of radical technological innovations (Hafeez, 2019).

Furthermore, the results indicated that entity theorists' evaluation of a social robot was higher when they thought of it as a servant. This finding appears consistent with that of Han et al. (2019) and lends partial support for Hypothesis 4, which stated in part, that relative to incremental theorists, entity theorists would prefer a robot positioned as a servant.

2.5.1 Limitations and Future Directions

At least three limitations of Study 1 are noteworthy. First, the robot role manipulation check was unsuccessful, and, in consequence, Hypothesis 4 could not be directly tested. It is unclear whether this was due to the manipulation, the measure, or both. One possible explanation is that there may have been an order effect, arising from the fact that the servant item in the manipulation check questionnaire, was presented first in both conditions. If this were the case, participants in the assistant condition may have mistakenly assumed that a robot positioned as a servant was equivalent to one positioned as an assistant.

A further possible explanation could be that the manipulation text was too subtle, or that some participants did not read it or only skimmed through, and were thus, not exposed to the experimental manipulation. The latter explanation is consistent with previous findings, which have demonstrated that inattention and low-effort are typical of Mturk samples to varying degrees (see Ford, 2017; Chandler et al., 2013; Hauser et al., 2019). Nevertheless, future studies should attempt to replicate and explore this effect further as it was unanticipated.

Additionally, there appears to be more incremental theorists in the sample than entity theorists (see Figure. 2.2). This observation is not entirely surprising, and has been noted by other researchers from different domains (see Heslin et al., 2005; Hughes, 2015; Schumann and Dweck, 2014; Iwai and de França Carvalho, 2020), suggesting that entity theorists might be less likely to participate in research studies than incremental theorists.

Another limitation of Study 1 lies in the fact that written descriptions were utilized. Even though this was a methodological decision, explained in Section 2.3.4, there would almost certainly have been some variability in the way participants imagined robot embodiment and appearance (Rueben et al., 2020), which may have influenced some participant's responses to the measures used. Accordingly, it is crucial for subsequent studies to examine implicit self-theory effects with real social robots, such as NAO (see Gouaillier et al., 2009), and with richer stimulus materials (e.g., videos, images, in-person human-robot interactions).

Finally, participants' implicit self-theory orientation was not manipulated. As discussed in Section 1.2, this procedure is used extensively to assess the causal role of implicit self-theories (see also Yorkston et al., 2010; Park and John, 2012). Further research is desirous in order to test the experimental manipulation of implicit self-theory in an HRI context.

2.6 Conclusion

The findings described in this chapter provide the first empirical support for the proposition that implicit self-theory influences, at least to some degree, one's perceptions and inferences about social robots. Specifically, these results demonstrate the influential role of implicit self-theories on one's robot anxiety and robot evaluation. The next chapter builds on, and substantially extends, Study 1 by reporting the results of Study 2, which used an in-person experimental design to examine the effects of implicit self-theories on people's responses to a robot that praised for ability (i.e., intelligence), and for effort (i.e., hard work), after completing a difficult reading and comprehension task.

Chapter III

Study 2

III

Study 2: The Influence of Implicit Self-Theories on People's Evaluations of a Robot After Perceived Success

Material covered in this Chapter has previously been published in Allan, D. D., Vonasch, A. J., and Bartneck, C. (2022c). "i have to praise you like i should?" the effects of implicit self-theories and robot-delivered praise on evaluations of a social robot. *International Journal of Social Robotics*

3.1 Introduction

This chapter presents the results of Study 2, which sought to, among other things, assess for the first time, the effect of implicit self-theories in an in-person human-robot interaction. More particularly, it was theorized that implicit self-theories and robot-delivered praise could interactively influence the way people evaluate a social robot, after apparently succeeding on a challenging task. The rest of this chapter is organised as follows: Section 3.2 provides the study overview, and Section 3.2.1 presents the hypotheses. Following this, Section 3.3. describes the method used, and Section 3.4 presents the experimental results. After this, Section 3.5 considers some inherent limitations of the study and identifies directions for future work. Section 3.6 provides a brief conclusion.

3.2 Overview

The rationale for Study 2 is underscored by the following research evidence. First, and as indicated in Chapter 1, the differences in implicit self-theory are most salient under conditions of challenge or when facing setbacks (Yeager and Dweck, 2020).

Second, ample research suggests that feedback and praise, particularly, effort (i.e., hard work) and ability (i.e., intelligence) praise, may foster implicit self-theories from an early age on, as noted in Chapter 1 (see Footnote 4).

It is important to point out that the study of how implicit self-theories are formed can be traced to Mueller and Dweck's (1998) landmark praise study.

In this study, fifth-grade children worked on a set of challenging tasks and were subsequently praised, first with outcome praise ("Wow, you did very well on these problems. You got [number of problems] right. That's a really high score.") The children were informed that of the problems they had answered they had solved at least 80%. This was followed by one of two types of praise—ability (e.g., "You must be smart at these problems."), and effort (e.g., "You must have worked hard at these problems.") A control condition received the initial outcome praise only. Next, the children were given a set of problems of increased difficulty, of which most failed. According to Mueller and Dweck (1998), the children who were praised for their ability viewed their intelligence as innate. Accordingly, they rejected a hard task, in favor of an easier task that, presumably, would pose no threat to their intelligence. Contrastingly, the children praised for their effort viewed their intelligence as something that could be developed. Consequently, this group chose a hard task rather than an easier task in order to learn from it.

Since this study, researchers have continued to find support for the link between ability and effort-focused praise and the development of one's implicit self-theory (e.g., Cimpian et al., 2007; Haimovitz and Henderlong Corpus, 2011; Kamins and Dweck, 1999; Pomerantz and Kempner, 2013; Skipper and Douglas, 2012; Lou and Noels, 2020). For example, Gunderson and colleagues examined parents' use of ability and effort praise in the presence of their children aged 1–3 years. They demonstrated that parents' use of effort praise predicted children's incremental theory of intelligence 5 years later (Gunderson et al., 2013). Furthermore, in a study conducted by Pomerantz and Kempner (2013), parents of 8 to 10 year-olds were interviewed on a daily basis, in order to track their use of ability and effort-focused praise. Upon assessing the children's implicit self-theories six months later, they found that frequent use of ability praise predicted children's entity self-theory.

In addition to its relevance to implicit self-theories, praise is also a topic of broad interest in the field of HRI. Praise has been conducted from different perspectives in HRI. Some work has focused on: investigating how humans attribute praise and punishment to robots (Bartneck et al., 2006, 2008), robot-delivered praise for increasing user motivation (Fasola and Matarić, 2013; Schneider and Kummert, 2016), self-efficacy (Zafari et al., 2019), as well as personalized robot praise toward children (Serholt and Barendregt, 2016) and older adults (Thompson et al., 2017); other work has evaluated the role of praise in game-based interactions (Ali et al., 2021) and nurturing praise in a therapy context (Tapus et al., 2008), and still other research has examined the relationship between robot-delivered praise, trust, and compliance (Ghazali et al., 2019).

Despite progress made, it is striking that there is scant scholarly research on the role of ability-focused or effort-focused praise in the field (but see Davison et al., 2021, discussed in Section 1.2). This is a surprising omission given that this form of praise is one of the most extensively researched praise types in the educational and psychological literatures (see Lam et al., 2008; Covington and Omelich, 1985; Brummelman, 2020), not to mention it has received attention from researchers in adjacent fields such as Human-Computer Interaction (e.g., Tzeng and Chen, 2012).

As such, it is unknown whether one's implicit self-theory orientation and robot delivered-praise can interactively determine one's evaluation of a robot. Accordingly, the study reported herein was designed to empirically investigate this issue.

3.2.1 Hypotheses

As detailed in Chapter 1, Section 1.2, entity theorists' basic assumption that traits are fixed leads them to devalue effort and constantly seek validation (Dweck and Leggett, 1988; Elliott and Dweck, 1988), praise (Murphy and Dweck, 2016b), and favorable judgments about their performance (Ommundsen et al., 2005). According to Dweck (2008) entity theorists' sense of worth rests on demonstrating these traits (see also Blackwell et al., 2007; Dweck, 2000). Relatedly, those endorsing an entity theory are also more likely to describe their ideal romantic partner as a person who would bolster their fixed qualities (i.e., praise them).

In contrast, incremental theorists prefer romantic partners that will encourage their development (see Dweck, 2008). Furthermore, incremental theorists are motivated by confronting a challenge, learning, and developing their ability (Molden and Dweck, 2006; Jain et al., 2009; Robins and Pals, 2002). Moreover, they are not generally content with gaining favorable competence feedback about themselves (Dweck and Leggett, 1988; Mathur et al., 2014). Hence, given that incremental theorists are inclined to value effort, and self-monitor their own progress (Dweck and Leggett, 1988; Mathur et al., 2014), external praise that provides flattering extrinsic validation of their capabilities is likely to carry less meaning for incremental theorists than entity theorists.

Finally, it seems pertinent to note that in Study 1 incremental theorists were inclined to evaluate robots favorably, regardless of how a robot was positioned (in this case, servant, or assistant).

Considering all the above, entity theorists would be expected to be more partial to a robot that praises for ability than one that praises for effort. Additionally, incremental theorists should be relatively unaffected by either effort or ability-focused praise, and, in turn, rate the robot favorably in both conditions. Stated formally:

Hypothesis 1 (H1): Entity theorists will evaluate the robot as more (versus less) likable after receiving ability (versus effort) praise.

Hypothesis 2 (H2): Incremental theorists will evaluate the robot as higher (versus lower) in likability after receiving both effort and ability praise.

Hypothesis 3 (H3): Entity theorists will evaluate the robot as more (versus less) intelligent after receiving ability (versus effort) praise.

Hypothesis 4 (H4): Incremental theorists will evaluate the robot as higher (versus lower) in intelligence after receiving both effort and ability praise.

3.3 Method

The preceding hypotheses were tested with an in-person laboratory experiment. Study 2 was performed in two steps. In the first step, participants' implicit self-theories of intelligence were manipulated impelling some toward an incremental theory and others toward an entity theory. During the second step, a difficult reading and comprehension task was administered, on which participants were told they did well. Participants were then presented with either effort or ability praise. Participants' likability and perceived intelligence scores were then measured. Participants were randomized into four conditions: (1) entity theory/effort praise, (2) entity theory/ability praise, (3) incremental theory/effort praise, and (4) incremental theory/ability praise. The experiment was conducted by one of two experimenters (both male) who were dressed professionally and wore the same clothing. Both experimenters followed a detailed script. See `https://osf.io/j4mvg/` for the full script.

The study hypotheses and analysis plan were pre-registered and can be found here: `https://osf.io/54nhk/`. The study took place in May 2021. Data collection took place over the course of five days.[1]

The experimental procedure was reviewed and approved by the Human Research Ethics Committee of the University of Canterbury (HEC 2020/130).

3.3.1 Recruitment

Participants were recruited via social media postings, recruiting websites; and numerous posters, flyer drops, and word of mouth around the University of Canterbury and the surrounding community. All advertisements indicated that the study was designed to investigate the suitability of a robot as a "test marker" (e.g., marking tests and exams in order to support educators) in the context of a reading and multiple-choice exercise. The robot test marker context allowed the theory inductions to be carried out, inconspicuously, as well as to give participants credible robot praise on a plausibly challenging task. Sample recruitment materials are available at `https://osf.io/z5x8s/` and `https://osf.io/tqxcp/`.

Inclusion criteria comprised (1) Participants 18 years or older with fluent English and basic reading and comprehension skills, and (2) no experience with robotics and/or any expertise in AI (e.g., data science and machine learning). Past research suggests that some forms of technical expertise and experience with robotics, both independently and in combination, have positive and moderating effects on people's attitudes toward, and acceptance of, social robots (see De Graaf and Allouch, 2013; Lee and Šabanović, 2014; Dudek et al., 2020; Latikka et al., 2021; Turja and Oksanen, 2019; Kühnlenz et al., 2018; Sanders et al., 2017). Therefore, it is reasonable to presume that the inclusion of such individuals may have potentially produced distorted findings. An a priori power analysis indicated that the sample size ($n = 120$) was adequate for a medium effect size (based on the Cohen criteria; Cohen, 1992) at 0.80 power and an alpha level of .05 to detect effects through analysis of variance (ANOVA). Accordingly, the pre-registered sample size was 120 participants, or as many participants as possible before May 14, 2021.

3.3.2 Design

Study 2 used a 2 (implicit self-theory: entity vs. incremental theory) $\times 2$ (ability vs. effort praise) between-subjects factorial design.

[1] It is important to make clear that in the pre-registration, it was declared that data collection process had been started at the time of preregistration. This, however, pertained only to the recruitment component. In other words, the research described herein had not, at that time been conducted in any way whatsoever.

Figure 3.1: The robot used in Study 2: A NAO V5 by SoftBank Robotics

3.3.3 Participants

A total of 101 adults participated for a $10 gift card. One participant was precluded from analyses a-priori on the basis of failing to answer the manipulation check correctly. Thus, the sample used for analyses comprised 100 participants (53 female, 44 male, 3 with no gender reported) from Christchurch, New Zealand. They ranged in age from 18 to 73 years old ($M = 28.4, SD = 15.3$). All participants had a high-school education, and 31% reported having either an undergraduate degree or a postgraduate degree. All participants gave informed consent prior to participating.

3.3.4 Procedure

Participants were tested individually in a small room with a round flat desk. Each participant was briefed about the nature of the study and introduced to the social robot (A humanoid NAO V5[2] from Softbank Robotics; see Figure. 3.1). The robot was seated on a table in front of the participant for the duration of the experiment (see Figure. 3.2).

After providing consent, and completing a brief demographics questionnaire (age, gender, education), participants were given a 1 page document that described the robot's supposed intelligent reasoning capabilities (a PDF version is available at `https://osf.io/zjvap/`). This aspect of the cover story was designed to ensure that participants perceived the robot praise as well-reasoned, rather than simply a pre-programmed response (which, in fact, it was).

Participants were then instructed to complete a pre-study task, in which their implicit self-theories were manipulated (described later in Section 3.3.5). Specifically, they were told they had six minutes to read a short article on an iPad, and two and a half minutes to complete some questions associated with that article. At

[2] `https://www.softbankrobotics.com/emea/en/nao`

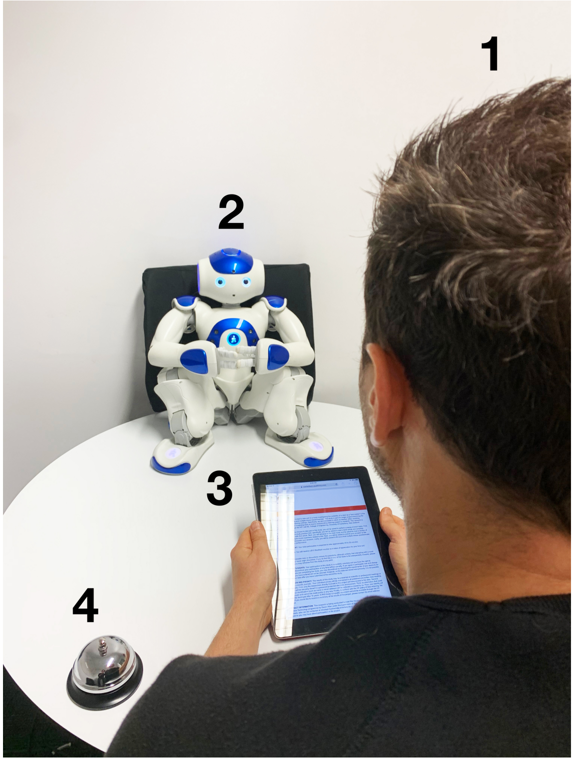


Figure 3.2: Experimental setup with the following components: 1) Participant, 2) NAO robot, 3) iPad, and 4) Bell.

this point, they were informed that the study was about assessing the robot not them. Further, they were told that their answers would be automatically converted to a special NAO code.[3] They were then informed that after they had answered the questions, they were to hold the iPad up in front of the robot's eyes and wait until the robot said, "OK got it".

When they had completed the pre-study task, participants were told that they would advance to the primary study task. They were subsequently advised that they would have three minutes to read a section of text and two minutes to answer some questions associated with that text (described later in Section 3.3.5). Moreover, they were told not to worry if the text didn't appear to make sense at first, as it had been selected purely because the content and the data produced, would be a suitable challenge for the robot to assess. Once again, participants were advised that the purpose of the study was to assess the robot, not them. The participant was then told that their answers would be converted to a NAO code, as in the pre-study task and they would be required to hold it up to the robot's eyes until it says, "OK got it". They were also advised that the experimenter would leave the room until the participant had finished, so as not to distract them. As well, the participant was instructed to ring the bell when the robot told them to do so. The experimenter then left the room.

After they had completed the study task, participants presented the NAO code to the robot. At this point, the robot ostensibly scanned the test and delivered the praise manipulation (outlined next in Section 3.3.5). After delivering this praise,

[3] A NAO code or "Naomark" is a circular symbol specific to the NAO robot. The robot is able to detect, and respond to this symbol (see `http://doc.aldebaran.com/2-1/naoqi/vision/allandmarkdetection.html`).

the robot thanked the participant for their participation and told them to ring the bell. Afterward, the experimenter entered the room and asked the robot how the test went, to which the robot replied with either effort or ability-related praise. The robot then asked if it should go into standby mode, the experimenter replied, "Yes", and then informed the participant that the robot could no longer see or hear them. This was done to encourage participants' candid feedback. Next, participants completed the measures described later in Section 3.3.6. After completing these items, participants were informed that the study was over, and debriefed. In keeping with the debriefing procedure used by Nussbaum and Dweck (2008), special care was taken to expound the fictitious nature of the theory induction article. Especially, the fact that intelligence is understood to have both malleable and stable qualities (Dweck, 2013). Finally, participants were compensated for their participation, informally probed for suspicion,[4] then dismissed.

3.3.5 Materials

Implicit Self-Theory Manipulation

Participants were randomly assigned to read one of two scientific *Psychology Today* articles adapted from Bergen (1991), endorsing either an incremental (e.g., "up to eighty-eight percent of a person's intelligence is due to environmental factors"), or entity (e.g., "up to eighty-eight percent of a person's intelligence is due to genetic factors") theory of intelligence. This priming stimuli is commonly used and well-validated in implicit self-theory research (see Bergen, 1991; Dinger and Dickhäuser, 2013; Hong et al., 1999; Schroder et al., 2014; Huang et al., 2017). After reading their respective articles, participants were asked to (a) summarize the main point of the article in no more than three sentences, and (b) complete the implicit theory of intelligence measure—which served as a manipulation check—described later in Section 3.3.6. It seems important to underscore that the theory manipulation was performed following established procedures.

Task Materials

A challenging reading test was set based on Nussbaum and Dweck (2008), for which praise was later given. The rationale for this was to provide a challenging stimulus because the effects of implicit self-theory are most pronounced under conditions of challenge or difficulty (Rattan et al., 2012), as discussed in Section 3.2. In particular, participants were given three minutes to read an excerpt from Freud's *The Interpretation of Dreams*. Following Nussbaum and Dweck the text was chosen to be reasonably perplexing, and the three minute time limit was purposely insufficient so as to allow for minimal comprehension. Furthermore, participants were instructed to answer five multiple-choice questions associated with the passage. To ensure the participants felt uncertain of their performance and thus found the robot-delivered praise to be somewhat credible, the questions and answers were designed to be vague and ambiguous.

[4]None of the participants reported suspicion about any aspect of the study.

Manipulation of Praise

Drawing, in part, on the procedure outlined by Mueller and Dweck (1998), the praise the robot administered was manipulated. Specifically, the robot informed all participants that they had performed well (e.g., "Wow, you did very well"). It is worthwhile emphasizing here, that all participants were explicitly told they had been successful in terms of their task performance. (It might be more descriptively accurate, in some sense, to say that the praise was delivered after a successful performance on a difficult task). Following the outcome praise, participants were told that they had 60% of the questions correct.[5] They subsequently heard one of two types of praise from the robot: some were praised for their effort (e.g., "You must have worked hard at these questions" and "Interesting, the pattern indicates that you tend to put in a lot of effort when faced with challenges, would that be right? Just answer yes or no"). Whereas others were praised for their ability (e.g., "You must be smart at these questions" and "Interesting, the pattern indicates that you tend to rely on your intelligence when faced with challenges, would that be right? Just answer yes or no"). Additionally, when the experimenter re-entered the room and asked the robot how the test went, the robot replied, "Overall, the participant has done well at these questions, although this participant is not one of the top performers, the pattern indicates they are quite a hard worker". Alternatively, the robot said, "Overall, the participant has done well at these questions, although this participant is not one of the best performers, the pattern indicates they are quite smart". It is important to point out that the justification for having the robot repeat the praise is based on prior work suggesting that robots need to deliver feedback multiple times in order for it to register as meaningful (Fasola and Matarić, 2013). See `https://osf.io/cs42y/` and `https://osf.io/yu37e/` for video examples of both conditions.

3.3.6 Measures

Manipulation Check

The six-item Implicit Theories of Intelligence Scale (ITIS; Dweck, 2000) was used as a manipulation check for participants' implicit self-theories. This is a well-established procedure for checking the effectiveness of primed implicit self-theories (see Levy and Dweck, 1998; Mathur et al., 2014; Rai and Lin, 2019). This scale has three items that measure incremental beliefs (e.g., "You have a certain amount of intelligence, and you really can't do much to change it") and three that measure entity theory beliefs (e.g., "You can always greatly change how intelligent you are"). Participants provided responses using a 6-point Likert scale (1=strongly disagree, 6=strongly agree). The incremental items were then reverse scored, and a mean score calculated for all six items, with high scores representing greater endorsement of an incremental theory of intelligence. The internal consistency of this measure was excellent ($\alpha = 0.95,\ M = 2.84,\ SD = 1.48$).

[5] 60% was used instead of 80% adopted by Mueller and Dweck (1998), because there were only 5 questions used in the study described herein. Therefore, it was reasoned that 60% (3/5) was more tenable than 80% (4/5) with respect to participants' estimation of their own performance.

Perceived Intelligence

Perceived intelligence was assessed using the five-item Perceived Intelligence sub-scale of the Godspeed Questionnaire Series (GQS; Bartneck et al., 2009). Responses were made using a 5-point scale (e.g., 1=Unintelligent, 5=Intelligent, and 1=Ignorant, 5=Knowledgeable). This scale has been used in previous research to measure people's impressions of robots post-interaction (e.g., Haring et al., 2015; Ghiglino et al., 2021; Chirapornchai et al., 2021), particularly NAO robots (see Weiss and Bartneck, 2015). This measure exhibited excellent internal consistency in this sample ($\alpha = 0.90,\ M = 3.97,\ SD = 0.95$).

Likability

Likability was assessed with the five-item Likability sub-scale of the GQS [6] (Bartneck et al., 2009). Participants provided responses using a 5-point scale (e.g., 1=Dislike, 5=Like, and 1=Unfriendly, 5=Friendly). The internal consistency of this measure was excellent ($\alpha = 0.93,\ M = 4.13,\ SD = 0.93$).

3.4 Results

The dataset for this study can be found at `https://osf.io/j69nv/`.

3.4.1 Data Analysis

Manipulation Check

An independent sample t-test on ITIS scores indicated that participants who received the entity theory article scored higher on entity theory ($M = 1.66,\ SD = 1.13$) than those who received the incremental theory article ($M = 3.93,\ SD = 0.729$), $t(98) = -12,\ p < .001\ \ d = -2.41$, (95% CI [-3.03, -1.78]), thus, our implicit self-theory manipulation was successful. It is worth noting again, that lower scores on the ITIS represent higher entity beliefs.

Perceived Intelligence

A 2 (implicit self-theory) $\times 2$ (praise) between-subjects ANOVA on perceived intelligence revealed a significant interaction between implicit self-theory and praise on perceived intelligence ($F(1, 96) = 24.0,\ p < .001$). Follow up planned comparisons showed that participants who were primed with entity theory demonstrated higher perceived intelligence scores for the robot when praised for ability

[6] One might reasonably wonder why only two sub-scales were used to measure robot evaluation. There were two reasons for this decision. First, these scales were the most pertinent to the hypotheses under investigation (see Section 3.2.1). Support for the use of some, but not all, of the GQS items, is offered by a qualitative meta-analysis which found that 62.5% of studies did not use the entire GQS questionnaire. Rather, there was a moderate preference for the use of the perceived intelligence and likeability scales (Weiss and Bartneck, 2015). At the same time, it must be acknowledged that the GQS is the most widely used instrument for assessing perceptions and evaluations of robots in HRI (see Zimmerman et al., 2022). Secondly, and perhaps most importantly, the use of two short items helped to reduce study fatigue on the part of participants. This was particularly necessary given the cognitively demanding nature of the experimental tasks.

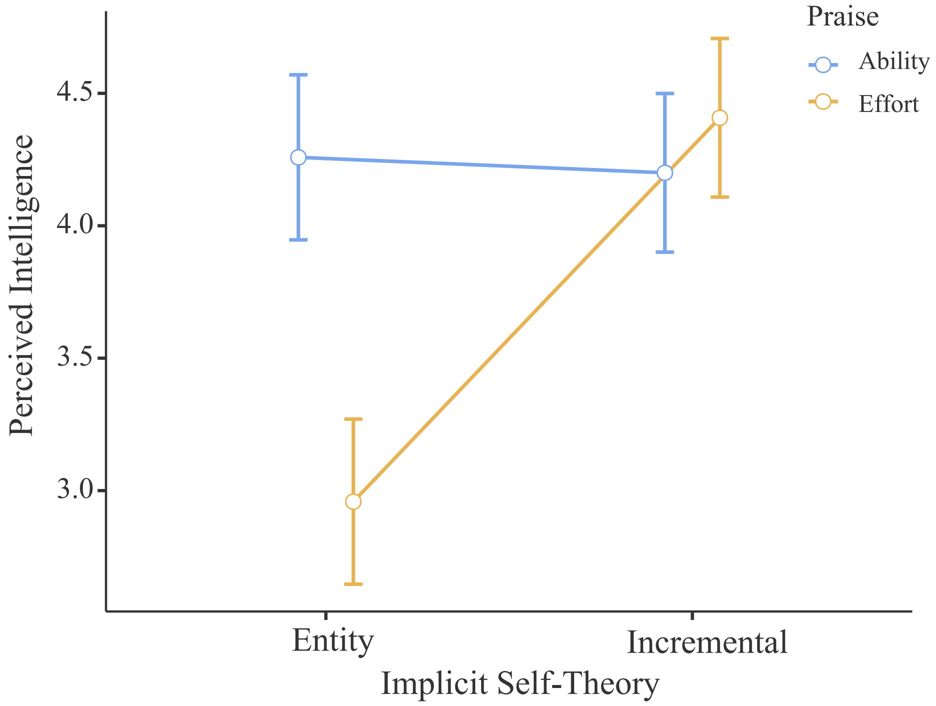


Figure 3.3: Perceived intelligence means and 95% confidence intervals for the interaction between implicit self-theories and praise.

($M = 4.26,\ SD = 0.606$) versus effort ($M = 2.96,\ SD = 1.26$), $t(96) = 5.855,\ p <$ $.001,\ d = 1.69$, (95% CI [1.06, 2.31]). However, results for participants manipulated to believe in incremental self-theory showed no significant difference between ability ($M = 4.20,\ SD = 0.534$) and effort-focused praise on ratings of perceived intelligence ($M = 4.41,\ SD = 0.447$), $t(96) = -.974,\ p = .765,\ d = -0.27$, (95% CI [-0.82, 0.28]). Given these results, hypotheses H3 and H4 were supported (see Figure. 3.3).

Likeability

A similar ANOVA test on participant's ratings of likability was also conducted. As expected, a significant interaction between implicit self-theory and praise on ratings of likability was observed ($F(1, 96) = 5.30,\ p = .024$). Planned comparison results revealed that for entity theorists, ability-focused praise led to significantly higher likability ratings ($M = 4.24,\ SD = 0.698$) than effort-focused praise ($M = 3.38,\ SD = 1.20$), $t(96) = 3.681,\ p = .002,\ d = 1.06$, (95% CI [0.47, -1.65]). Furthermore, no significant difference between ability ($M = 4.48,\ SD = 0.552$) and effort-focused praise on ratings of likability ($M = 4.37,\ SD = 0.694$), $t(96) = -.510,\ p = .956,\ d = 0.14$, (95% CI [-0.40, 0.69]), was observed for incremental theorists. These findings supported H1 and H2 (see Figure. 3.4).

3.5 Discussion

The findings presented herein provide evidence that implicit self-theories (entity vs. incremental), and robot-delivered praise (ability vs. effort) can interactively influ-

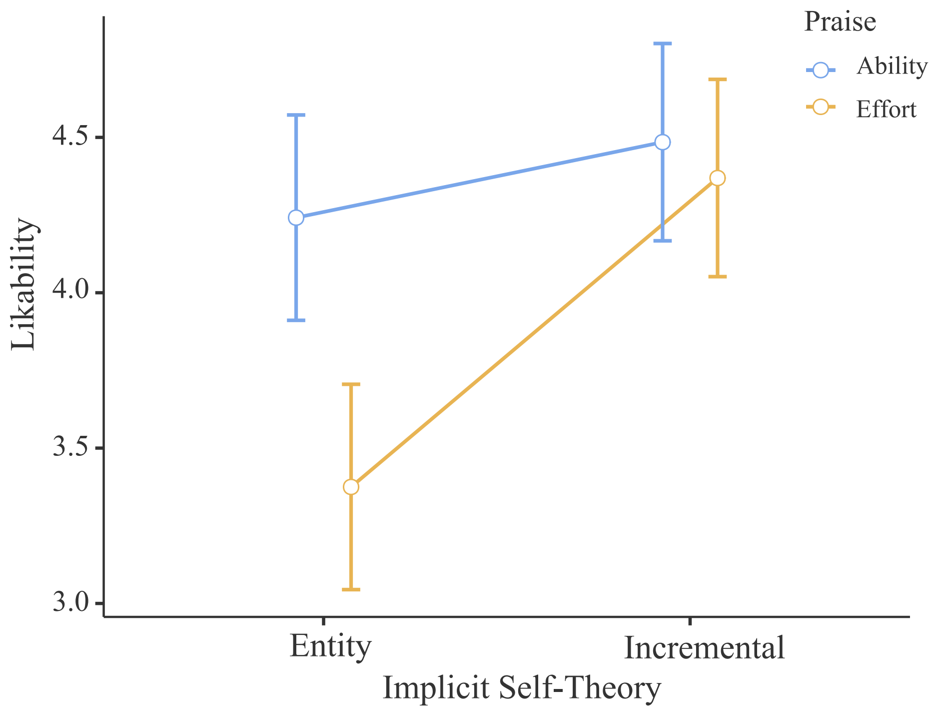


Figure 3.4: Likability means and 95% confidence intervals for the interaction between implicit self-theories and praise.

ence the way people evaluate social robots after a challenging task. More precisely, Study 2 demonstrates that entity theorists, who are prone to conceal their shortcomings and seek out favorable judgments from others, rate a robot lower in both perceived intelligence (H1) and likability (H2), when it praises them for effort (vs. ability).

Conversely, incremental theorists, who are more self-driven and learning-oriented, appear unaffected by praise type, such that they evaluate a robot high in both perceived intelligence (H3) and likability (H4), regardless of the praise it delivers.

On top of all that, Study 2 addressed two limitations of the first study. In Study 1, participants read about social robots, they did not view or interact with any. An abundance of HRI studies show that exposure to a physically present robot is both more preferable and more accurate in eliciting people's direct responses and evaluations of robots (see Haring et al., 2014; Horstmann et al., 2018; Savela et al., 2018). In consequence, Study 2 examined, for the first time, the role of implicit self-theories in an in-person human-robot interaction.

And, while in Study 1, participants' implicit self-theories were measured as an individual difference variable, in Study 2, participants' implicit self-theories were directly manipulated, providing the first evidence of the experimental manipulation of implicit self-theories in HRI.

3.5.1 Limitations and Future Directions

Though strong support was found for the study hypotheses, these findings were nonetheless subject to certain limitations. First, generalizations from these findings may be limited by the decision to use the NAO V5 robot. However, it might be mentioned here that some, such as Keizer and colleagues (2019), have argued that the NAO does not differ meaningfully from current-day humanoid robots in terms of features, capabilities, and intended use. Nonetheless, future work employing alternative humanoid robots (e.g., Pepper, Zeno, or even Baxter) is warranted.

Beyond investigating perceived intelligence and likability, it will also be important for future research to test how other dimensions of robot evaluation relate with praise type and implicit self-theory. For instance, Spatola and colleagues (2021) have explored the association of perceived competency and warmth on evaluations of social robots. This, they argued, forms a type of social evaluation, such that social robots perceived as warm (vs. cold) and competent (vs. incompetent) are viewed more (vs. less) favourably. Therefore, future research could replicate this study using alternative evaluation frameworks, such as items from the Robotic Social Attribute Scale (RoSAS; Carpinella et al., 2017) and the Human–Robot Interaction Evaluation Scale (HRIES; Spatola et al., 2021).

Another potential limitation might be that a control condition was not included in the study design. Although, this decision was made on the basis of time, space, and resource constraints (i.e., the inclusion of a control condition would have required 50% more participants), subsequent studies should include a control condition to compare with the experimental group.

3.6 Conclusion

The findings of Study 2 provide further evidence for the proposition that implicit self-theory is an influential variable that determines how individuals perceive of, evaluate, and respond to, social robots. Specifically, Study 2 demonstrates the first experimental evidence that entity theorists, as compared to incremental theorists, evaluate a social robot as less likable, and less intelligent, after receiving effort (versus ability) praise delivered after a successful performance on a difficult task. Whereas incremental theorists remain largely unaffected by either praise type and instead evaluate a social robot favorably regardless of the praise it delivers. The next chapter reports the results of Study 3, which adds to and further complements the results of Studies 1 and 2, by investigating the role of implicit self-theories on people's responses to a robot after experiencing perceived failure.

Chapter IV

Study 3

IV

Study 3: The Influence of Implicit Self-Theories on People's Responses to a Robot After Perceived Failure

Material covered in this Chapter has previously been published in Allan, D. D., Vonasch, A. J., and Bartneck, C. (2022a). Better than us: The role of implicit self-theories in determining perceived threat responses in hri. In *Proceedings of the 2022 ACM/IEEE International Conference on Human-Robot Interaction*, pages 215–224

4.1 Introduction

This chapter reports the findings of Study 3, which builds from the results of Studies 1 and 2 to investigate the effects of implicit self-theory on people's responses to an apparently autonomous intelligent robot, that defeats human quiz players in a general knowledge game. This chapter is organized as follows: In Section 4.2, the study overview and hypotheses are described, and Section 4.3 outlines the method used. This is followed by a discussion in Section 4.5, highlighting key limitations and suggestions for future work. Lastly, Section 4.6 serves as a conclusion.

4.2 Overview and Hypotheses

The results of Study 2 demonstrate that theory relevant praise after success on a challenging task, leads to favorable evaluations of a social robot. This leaves open the equally important and interesting question: What are the differential effects of implicit self-theory orientation on people's responses to a robot after they experience failure? Accordingly, Study 3 aimed to pursue this question.

More specifically, in Study 3, participants watched a video of an apparently autonomous intelligent robot defeating human quiz players in a general knowledge game. Following this, participants were asked to think about losing to the robot just as those in the video had. They subsequently received either social comparison feedback, improvement-oriented feedback, or no feedback, and were then given the

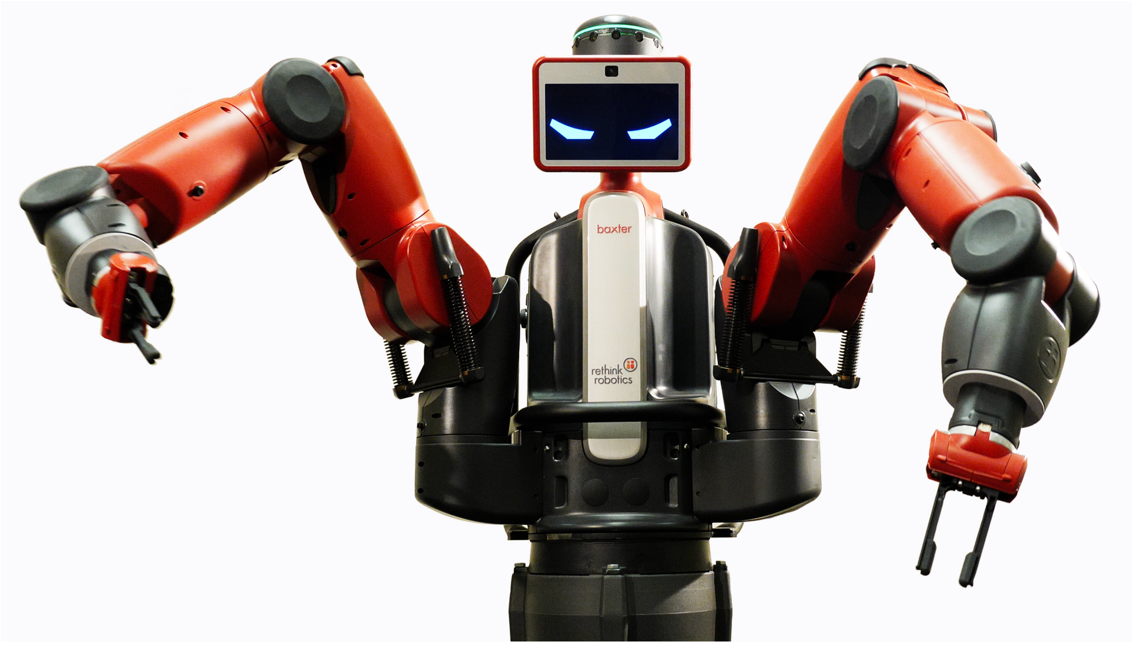


Figure 4.1: The Baxter robot used in Study 3.

opportunity to play against the robot. The rest of this section outlines the rationale underscoring this study.

As detailed in Chapter 1, Section 1.2, the central tenant of implicit self-theories is that they compel people toward, "different goals and attributions in situations involving challenges and setbacks" (Yeager and Dweck, 2020, p. 1270). For example, Nussbaum and Dweck (2008) showed that entity and incremental theorists exhibit distinct patterns of response following failure. In one set of studies, college students who worked on a difficult task on which they inevitably failed, were given the option to examine the strategies of others who had performed either better, worse, or the same as they had (i.e., upward or downward social comparisons). Entity participants opted to view the strategies of those who had performed comparatively worse than they had. Incremental participants, however, viewed this as an opportunity to redress their performance by choosing to look at the strategies of people who had exceeded their performance (Nussbaum and Dweck, 2008, Studies 1 and 2). A follow-up study confirmed that entity theorists felt better after defensively comparing themselves to poorer performers. Whereas incremental theorists felt better confronting and addressing their poor performance in an effort to improve on the task (Nussbaum and Dweck, 2008, Study 3).

As Dweck (2017) points out, when entity theorists, "opt for success over growth, what are they really trying to prove? That they're special. Even superior" (p. 29). Indeed, Study 1 provided initial evidence that entity theorists preferred a robot positioned as a servant, presumably, as a result of this effect (see Section 2.5).

To be sure, these findings are consistent with prior work showing that entity theorist are often driven by the aim of besting others in order to prove their competence (Dweck, 2009; Mangels et al., 2006), and are more likely to be intolerant of mistakes or setbacks (Dweck, 2013; Hong et al., 1999), whereas incremental theorists seek ways to increase competence through effort (Dweck and Leggett, 1988; Elliott and Dweck, 1988) and appear to be better able to recover after experiencing a setback (Elliott and Dweck, 1988; Rhodewalt, 1994; Robins and Pals, 2002).

Drawing on this nascent literature, it was expected that in comparison to incremental theorists, entity theorists with their desire to not appear incompetent, would be adverse to the idea of playing against a winning intelligent robot. It was also expected that entity theorists would evaluate a robot more favorably in the presence of social comparison feedback. Stated formally:

Hypothesis 1 (H1): Entity (vs. incremental) theorists will be less (vs. more) likely to indicate an interest in playing against the robot after imagining losing to it.

Hypothesis 2 (H2): Entity theorists will evaluate the robot in the social comparison condition more (vs. less) favorably than the robot in the improvement-oriented and control conditions.

As noted in Chapters 1 and 3 a substantial stream of implicit self-theory research has found that implicit self-theories can arise from, and be reinforced by, significant others' (e.g., teachers, parents, and peers) feedback (Haimovitz and Dweck, 2017). In one study (Lou and Noels, 2020), for example, university students with English as a second language failed a challenging English test and were given either ability-consoling feedback (e.g., "I'm sure you have great talent in other subjects."), improvement-oriented feedback (e.g., "If you put in the work, you'll be at the level of proficiency that you want."), or no feedback from a teacher. Students who received ability-consoling feedback reasoned that the teacher did not think they could improve and, in consequence, expressed an unwillingness to retake the test. Conversely, students who received the improvement-oriented feedback perceived the teacher to believe in their potential. According to Lou and Noels, the teacher's feedback strengthened and further increased students' endorsement of an incremental theory. Consistent with this notion, Study 2 found that implicit self-theories and robot-delivered feedback (praise, in this case), interactively influenced people's evaluations of a social robot after a challenging task. Together, these findings provide evidence of a matching effect between feedback and individual implicit self-theories (see also Cimpian et al., 2007; Haimovitz and Henderlong Corpus, 2011; Kamins and Dweck, 1999; Pomerantz and Kempner, 2013; Skipper and Douglas, 2012). Accordingly, it was expected that incremental theorists, with their focus on developing ability, would evaluate a robot more favorably in the presence of feedback that emphasizes improvement. Stated formally:

Hypothesis 3 (H3): Incremental theorists will evaluate the robot in the improvement-oriented condition more (vs. less) favorably than the robot in the social comparison feedback and control conditions.

Another frequently occurring finding in the implicit self-theory literature, is that an individual's implicit self-theory appears to influence the attributions they make regarding other people (Dweck et al., 1995a, 1993; Erdley and Dweck, 1993). For instance, Levy et al. (1998) found that, compared to incremental theorists, entity theorists endorsed extreme judgments on global traits (i.e., bad-good, evil-virtuous) toward a range of target groups (e.g., ethnic, occupational, and even fictitious groups). Moreover, those endorsing an entity theory have been found to display a significant increase in self-esteem when promoting stereotypes (Dweck, 2017).

Somewhat relatedly, converging research in HRI and HCI suggests that intelligent and highly capable robots, simultaneously threaten human safety and resources (i.e., realistic threat), as well as human uniqueness and identity (i.e., identity threat). This phenomenon is known as "perceived threat" in the literature.[1]

Yogeeswaran and colleagues (2016) for example, found, that a new generation of robots perceived to outperform humans on a variety of physical and mental tasks increased feelings of threat. This, according to the authors, is because human beings view such robots as members of a highly competent outgroup. Hence, these robots are considered to be plausible competition for resources and jobs, and, in addition, a threat to human identity and distinctiveness.

Correspondingly, in a study conducted by Złotowski et al. (2017), participants who watched videos of purportedly autonomous robots (i.e., robots capable of disregarding human commands), considered robots to be significantly more threatening (both in terms of realistic and identity threats) than those who were shown a video of what appeared to be non-autonomous robots. What this finding indicates, the authors argue, is that when human beings experience a sense of perceived loss and control over robots, such robots are perceived as threatening.

Other avenues of research have largely buttressed these findings (e.g., Strait et al., 2017; Hover et al., 2021; Stapels and Eyssel, 2021; Huang et al., 2021), particularly, Cha and his colleagues (2020) have made a strong case that human-machine intellectual comparisons induce perceived threats. In one study, they showed that the defeat of human Go champion, Lee Sedol by Google's AlphaGo computer program threatened human distinctiveness (Cha et al., 2020, Study 1). They also discovered that human intellectual comparison with machine intelligence adversely affects people's motivation towards intellectual tasks (Cha et al., 2020).

This finding is similar to that of Oh et al. (2017), who found that participants who viewed the Google DeepMind Challenge (AlphaGo vs. Lee Sedol), experienced not only threats to human distinctiveness and resources, but increased feelings of helplessness.

Notably, Study 1 demonstrated that entity theorists, relative to incremental theorists, exhibit greater negative robot beliefs. Interestingly, Złotowski et al. (2017), has demonstrated an association between negative robot beliefs and perceived threat. Therefore, the following hypotheses were formulated:

Hypothesis 4 (H4): Incremental (vs. entity) theorists will rate robots in general as posing less (vs. more) identity relevant threats.

Hypothesis 5 (H5): Entity (vs. incremental) theorists will rate robots in general as posing more (vs. less) realistic threats.

4.3 Method

The hypotheses were tested in an online human-subject experiment, which was administered through the Qualtrics online platform. Participants were drawn from

[1]It should be noted that perceived threat also known as "threat perception," is a topic of increasing import in the field of HRI (see for example Ferrari et al., 2016; Yogeeswaran et al., 2016; Złotowski et al., 2017; Meissner et al., 2020; Kieslich et al., 2021; Huang et al., 2021; Hover et al., 2021) in much the same way as praise (explored in Study 2), and robot acceptance (explored in Study 1) are.

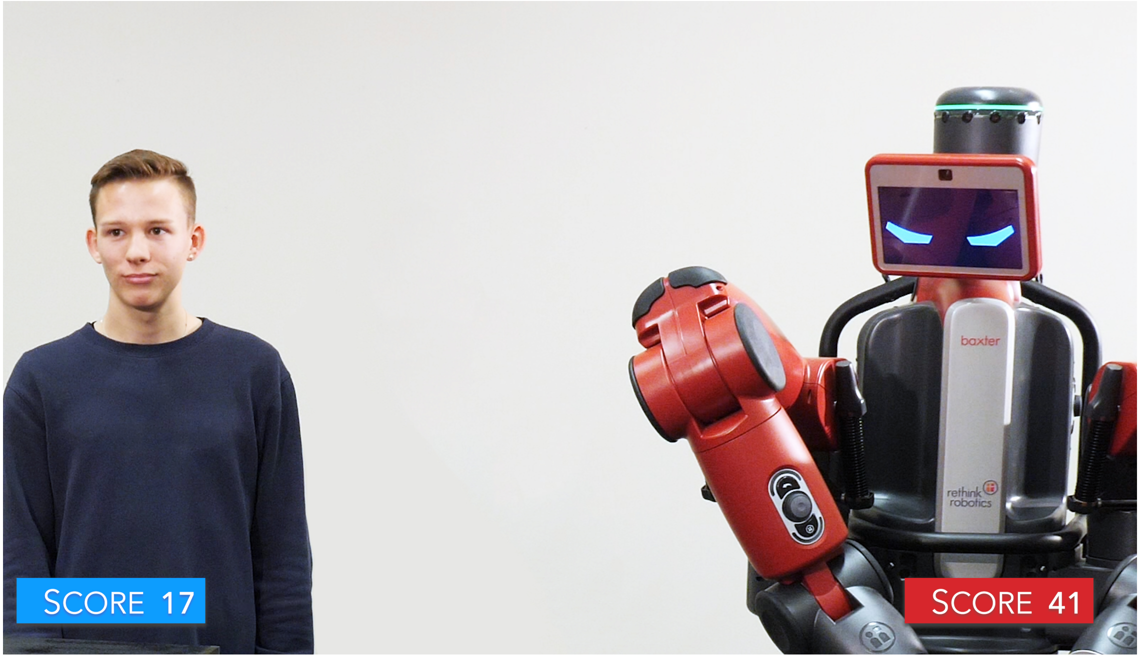


Figure 4.2: An on-the-spot screenshot from the video showing the Baxter robot and a human player. Note. The individual pictured here has provided written informed consent to publish their image alongside the thesis.

MTurk (described in Chapter 1, Section 2.3.1). While conducting online studies of this nature, is well accepted in HRI research (Hoffman and Zhao, 2020), it was particularly encouraged during the COVID-19 pandemic when Study 3 was carried out (see for example Feil-Seifer et al., 2020).

The experiment was performed in four parts. First, participants' implicit self-theories were experimentally primed. The main justification for inducing implicit self-theories, as opposed to measuring them as chronic orientations, was based on the findings of Study 1, which showed that MTurk workers who self-selected to take part in a study about robots, were prominently skewed in the direction of the incremental theory (See Figure. 2.2). Correspondingly, Study 3 sought to obtain a more theory-balanced sample. As in Study 2, Study 3 followed established implicit self-theory priming procedures (e.g., Bergen, 1991; Hong et al., 1997).

In the second part of the experiment, participants were exposed to a video of people playing against (and losing to) an ostensibly autonomous intelligent robot in a general knowledge quiz game (see Figure. 4.2). It is worth mentioning here, that the use of video media featuring robots has been shown repeatedly to be effective in engendering people's experiences of, and attitudes towards, autonomous robots (e.g., Bartneck and Keijsers, 2020), and most relevantly, perceived threats (see Yogeeswaran et al., 2016; Złotowski et al., 2017).

In the third part of the experiment, participants were presented with either improvement-oriented, social comparison, or neutral feedback. As noted in Section 4.2, such feedback has consistently emerged as an interactive influence with implicit self-theories on both people's responses to perceived failure (Mueller and Dweck, 1998), and judgments of others (Lou and Noels, 2020).

The fourth part of the experiment involved administering the dependent measures (described later in Section 4.3.7).

The experimental hypotheses and analysis plan were pre-registered at `https://`

`aspredicted.org/blind.php?x=/HY4_JJY`. All data were collected during August 2021, and the experimental protocol was reviewed and approved by the Human Research Ethics Committee of the University of Canterbury (HEC 2021/33).

4.3.1 Pretest

In light of recent evidence suggesting that an increasing number of participants on crowdsourcing websites are non-naïve (i.e., pre-exposed to a vast array of experimental materials and methods employed in the social sciences; Haug, 2018; Meyers et al., 2020), it was suspected that participants on MTurk might be familiar with extant (and increasingly distributed), implicit self-theory priming articles (e.g., Septianto et al., 2021; Aljukhadar and Senecal, 2021; Seo et al., 2021). Therefore, new manipulation articles were developed for Study 3. These articles were based on prior work (i.e., Chiu et al., 1997a; Bergen, 1991), but with specific focus on relevance and applicability in an HRI context.

The key excerpts for the entity theory article reads as follows:

> In his presentation Dr Silver argued that: "For most of us by the age of three, our intelligence level has set like plaster and will never soften again," he said. Dr. Silver reported on numerous large and compelling longitudinal studies which show people cannot increase their intelligence. He also reported on research that showed that up to 88 percent of a person's intelligence level is due to genetic factors, he explained: "About ten percent of intelligence seems to be determined during the first three years of life. This means that intelligence may be increased or decreased by only about two percent during most of a person's life." For Dr. Silver, the scientific consensus that human intelligence is largely fixed and unable to be changed is the catalyst for developing 'better-than-human' intelligent robots.

The key excerpt for the incremental theory article reads as follows:

> In his presentation Dr Silver argued that: "No one's intelligence is hard like a rock that it cannot be increased; only for some, greater effort and stimulation are needed to effect growth," he said. Dr. Silver reported on numerous large and compelling longitudinal studies which show that people can increase their intelligence. He also reported on research that showed that up to 88 percent of a person's intelligence level is due to environmental factors, he explained: "People may be born with a given level of intelligence, but we see increases in IQs of up to 50 points when people enter stimulating environments." For Dr. Silver, the scientific consensus that human intelligence is largely malleable and able to be changed is the catalyst for developing 'better-than-human' intelligent robots.

A pretest was then conducted in order to evaluate the effectiveness of the stimuli. Specifically (n=27 MTurk participants)[2] were randomly assigned to read either

[2]Sample size was based on the convention that a sample of 10 to 20 respondents is suitable for pretesting purposes (see Ricci-Cabello et al., 2019; Johanson and Brooks, 2010).

the incremental (n=15) or the entity article (n=12). To strengthen the manipulation, participants were then asked to summarize, in their own words the scientific consensus on intelligence. Additionally, participants were asked to indicate what Dr. Silver's research goal was. Next, participants completed The Implicit Theories of Intelligence Scale (ITIS; Dweck, 2000) which measured their primed implicit self-theory (described later in Section 4.3.7).

An independent sample t-test on the ITIS revealed that participants who received the entity theory article scored higher on entity theory beliefs ($M = 3.85,\ SD = 1.32$) than those who read the incremental theory article ($M = 4.81,\ SD = 1.03$), $t(25.0) = 2.13,\ p < .043\ \ d = -0.825$, (95% CI [-0.00, -1.63]). Thus, based on these results, the implicit self-theory prime was successful.[3] The full manipulation articles are available at `https://osf.io/t7f2j/`.

4.3.2 Design

Study 3 used a 2 (implicit self-theory: entity vs. incremental) ×3 (feedback: social comparison vs. improvement-oriented vs. control) between-subjects factorial design.

4.3.3 Recruitment

Participants were recruited through a task posted on the MTurk website. The posted task made clear that this was a study designed to survey people's "impressions of an intelligent autonomous game-playing robot." As well, prospective participants were informed that they would be required to (1) read a short paragraph describing "our approach to intelligent robots" and, (2) complete some questions associated with that paragraph. Furthermore, respondents were informed that they would be expected to watch a video of the robot and answer questions related to that video. This pretense provided the opportunity to unobtrusively administer the theory inductions, as well as present a plausible context in which an intelligent robot could convincingly defeat human beings. Participation was contingent on (a) having completed >50 surveys with a minimal HIT approval rate of 98% or greater, (b) being located in either the USA, UK, Canada, Ireland, Australia, or New Zealand, and (c) having no experience with robotics and/or any expertise in AI (the justification for this criteria has been discussed in Chapter 2, Section 3.3.1). An a priori power analysis conducted in G*Power (Faul et al., 2007) for a 2 ×3 ANOVA design, indicated that the sample size (n = 339) was appropriately necessary to detect a medium effect size (based on Cohen; Cohen, 1992) at 0.80 power and an alpha level of .05.

4.3.4 Participants

A total of 356 participants[4] recruited through MTurk agreed to participate in exchange for $1.00 in MTurk credit. Fifty participants (14.08%) who either answered the manipulation check question incorrectly (n=34), failed the attention

[3] As noted in previous chapters (and is perfectly worthwhile noting again), lower (vs. higher) scores indicate more (vs. less) of an entity theory.

[4] In order to account for probable exclusions, an additional 17 participants were recruited beyond the calculated sample size. This extended sample size was determined by budget constraints.

check (n=11), or who asked for their data to be excluded (n=6), were omitted from the dataset prior to analyses. The final sample consisted of 305 participants (173 male, 130 female, 2 with no gender reported), aged 18–29 (n=54), 30–39 (n=114), 40–49 (n=69), 50–59 (n=41), and 60 and over (n=27). Most participants (56.1%) ranked undergraduate education (some college education) as their highest level of educational attainment. While 26.2% held postgraduate degrees and 17.7% had a high school education alone. All participants provided consent prior to participation.

4.3.5 Procedure

Upon consent, participants were assigned randomly to read one of two articles, endorsing either an entity, or incremental theory of intelligence (described above in Section 4.3.1). Afterwards, participants were asked to briefly summarize their respective article, and complete The Implicit Theories of Intelligence Scale (ITIS; Dweck, 2000), which served as a manipulation check (described later in Section 4.3.7). Participants were then instructed to watch a video of a general knowledge quiz match, in which human quiz players appeared to compete against an apparently intelligent and autonomous robot (described below in Section 4.3.6). After viewing the video, participants were randomly assigned to one of three feedback conditions: improvement-oriented feedback, social comparison feedback, or neutral feedback (described below in Section 4.3.6). Following the manipulation, participants completed the dependent measures (described in Section 4.3.7), and provided their demographic details. Next, participants were debriefed about the true objective and the experimental rationale of the experiment, including the deception employed. Finally, participants were given an opportunity to withdraw their data, thanked, and compensated with $1.00.

4.3.6 Materials

Video

The video was designed to fit the tradition of real-world human-machine competitions, such as IBM Watson's Jeopardy! Challenge (Chandrasekar, 2014), and the Google DeepMind Challenge Match between Lee Sedol and AlphaGo (Lee et al., 2016). The video lasted approximately seven and a half minutes and featured a Baxter Robot (Guizzo and Ackerman, 2012). Baxter is a six foot three inch tall, humanoid robot, built initially for the manufacturing sector (see Figure. 4.1). Consistent with prior work (Nash et al., 2018), custom eyes with a random eye blink were designed principally for the study. In addition, the robot was pre-programmed with scripted responses, prior to filming. A female American voice introduced the robot, and following (Yogeeswaran et al., 2016), emphasized that Baxter had been "shown to outperform humans on both physical tasks, such as weight lifting and on mental tasks, such as chess and problem solving." Participants were then informed that what followed were "highlights of an intelligence challenge that took place between a Baxter Robot and players from a top university quiz team" (who were, in fact, actors performing scripted responses). Participants then watched as Baxter defeated the human players 4–0 in a best-of-four quiz game (see Figure. 4.3 for final scores). The quiz questions and answers were, in part, derived from the

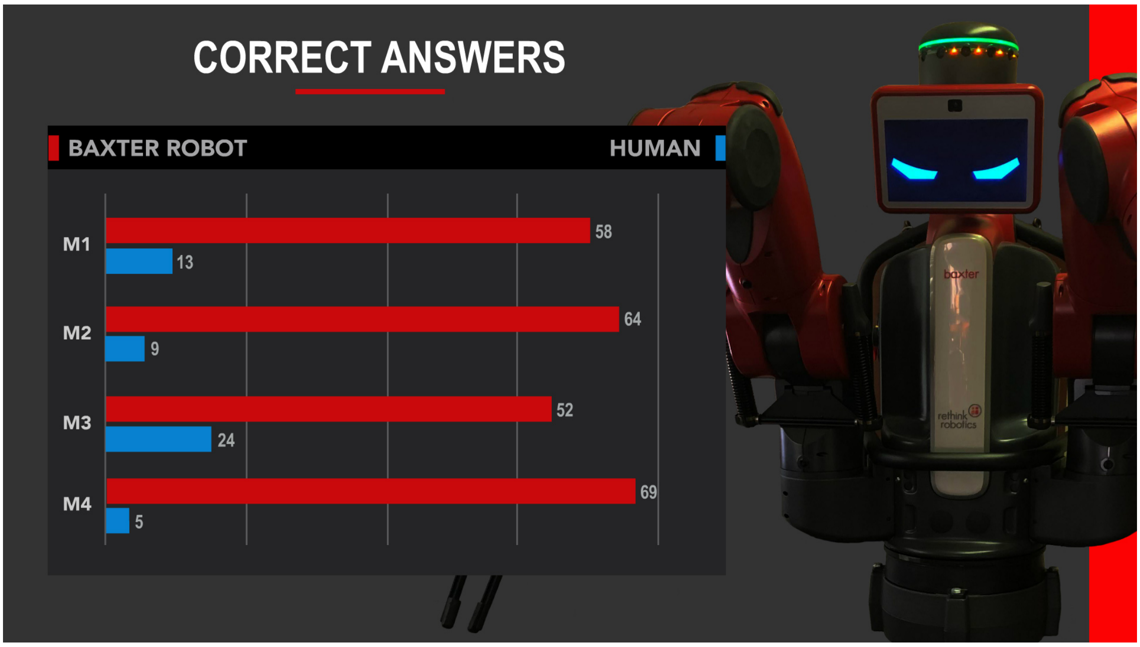


Figure 4.3: A screenshot of the final scores of the game.

official YouTube channel of the British television quiz show The Chase.[5] The video can be viewed online `https://youtu.be/VYPr-XUibFA`

Feedback

On the basis of prior work in implicit self-theories (discussed in Section 4.2), feedback was manipulated by emphasizing two different feedback types: improvement-oriented, and social comparison, which have observed interactive matching effects with individual self-theories (i.e., improvement-oriented/incremental, social comparison/entity), After watching the video, participants were instructed to imagine that they had lost to the robot, in the same way the players in the video had. Participants then read that the robot had provided feedback. Following (Lou and Noels, 2020), all participants were told that they had not performed well: "You did not do well in the game." Subsequently, some received improvement-oriented feedback (e.g., "Like with many things, practice makes perfect. If you put in the work, you'll surely improve. Do you want to play me again?"). Whereas others were exposed to social comparison feedback (i.e., "Most humans are not naturally good at general knowledge. However, you did 37% better than the other humans. Do you want to play me again?"). It might be useful to point out that 37% was derived from (Nussbaum and Dweck, 2008), with the aim of provoking the entity theorists, into engaging in downward social comparisons (i.e., to think of the people in the video who had done relatively poorly), and thus defensively repair their self-esteem. Finally, the control condition received no additional feedback ("Do you want to play me again?"). The feedback scripts were adapted somewhat from (Lou and Noels, 2020) to suit the study context.

[5] `https://www.youtube.com/c/thechaseofficial/videos`

4.3.7 Measures

Manipulation Check

The Implicit Theories of Intelligence Scale (ITIS; Dweck, 2000) was used as a manipulation check for participants' implicit self-theories (described in Chapter 2, Section 3.3.6). This measure demonstrated good internal consistency ($\alpha = 0.89$, $M = 3.60$, $SD = 1.22$).

Future Task Avoidance

A 5-item measure modified from (Lou and Noels, 2020) assessed participants' likelihood of playing against the robot after imagining losing to it (e.g., "I would try to avoid playing the robot again"). Participants indicated their agreement using a 5-point Likert scale (1=not at all, 5=very much). This measure demonstrated good internal consistency in this sample ($\alpha = 0.84$, $M = 3.23$, $SD = 1.08$).

Robot Evaluation

Robot Evaluation was assessed using the Robot Evaluation scale (described in Chapter 1, Section 2.3.5). The internal consistency of this measure was excellent ($\alpha = 0.94$, $M = 4.45$, $SD = 1.79$).

Identity Threat

Identity threat was assessed with the 5-item Identity Threat Measure (Yogeeswaran et al., 2016). Respondents indicated agreement with items (e.g., "Technological advancements in the area of robotics is threatening to human uniqueness") on a 7-point scale (1=strongly disagree, 7=strongly agree). This measure demonstrated excellent internal consistency in this sample ($\alpha = 0.93$, $M = 4.14$, $SD = 1.68$).

Realistic Threat

Realistic threat was assessed with the 5-item Realistic Threat Measure (Yogeeswaran et al., 2016), Agreement with items (e.g., "in the long run, robots pose a direct threat to human safety and wellbeing") was indicated on a 7-point scale (1-strongly disagree, 7=strongly agree). This measure demonstrated good internal consistency ($\alpha = 0.80$, $M = 4.47$, $SD = 1.28$).

4.4 Results

The dataset for Study 3 is publicly available at the OSF: `https://osf.io/rujge/`

4.4.1 Main Analyses

Manipulation Check

Results of an independent sample t-test verified that participants who read the entity theory article reported significantly higher ($t(303) = 9.18, p < .001$ $d = -1.05$, (95% CI [-1.05, -0.79]) entity theory beliefs ($M = 3.03$) than those who read the

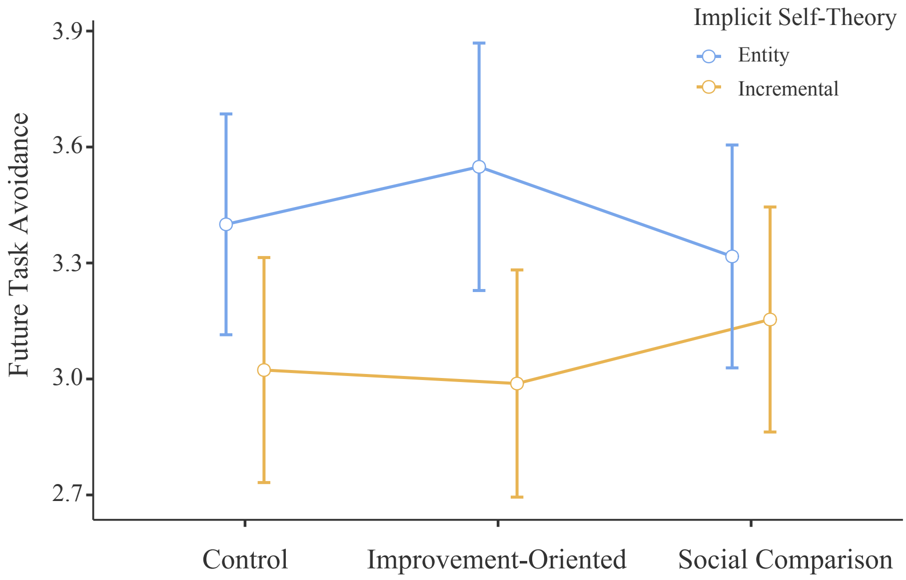


Figure 4.4: Future Task Avoidance means and 95% confidence intervals for the effects of implicit self-theories and feedback.

incremental theory article ($M = 4.16$); thus, the manipulation of implicit self-theory was successful.

Future Task Avoidance (H1)

A 2 (implicit self-theory) $\times 3$ (feedback) between-subjects ANOVA on future task performance was conducted. As expected, results showed a significant main effect for implicit self-theory ($F(1, 299) = 8.97,\ p = .003$). There were no predictions made regarding the effect of feedback condition, nor was there an interaction predicted. No significant effect of feedback condition, ($F(2, 299) = 0.07,\ p = .931$), and no significant interaction emerged ($F(2, 299) = 0.86,\ p = .423$). Planned comparisons indicated that participants who were primed with entity theory demonstrated higher task avoidance ($M = 3.42,\ SD = 1.00$) than those primed with incremental theory ($M = 3.06,\ SD = 1.12,\ t(299) = 3.00,\ p = .003,\ d = 0.344$, (95% CI [0.11, 0.57]). Consequently, H1 was supported (see Figure. 4.4).

Robot Evaluation (H2, H3)

An additional 2 (implicit self-theory) $\times 3$ (feedback) between-subjects ANOVA was performed to assess robot evaluation. Results revealed a significant main effect of implicit self-theory ($F(1, 299) = 62.87,\ p < .001$), but no main effect of feedback condition, ($F(2, 299) = 0.50,\ p = .606$). However, a significant interaction between implicit self-theory and feedback condition was observed ($F(2, 299) = 4.25,\ p = .015$). Planned comparisons were subsequently conducted to examine the hypothesized differences between conditions. As expected, planned comparison results for entity theorists revealed that robot evaluation was significantly higher in the social comparison condition ($M = 4.15,\ SD = 1.50$) than in the improvement-oriented condition ($M = 3.32,\ SD = 1.93$), $t(299) = -2.50,\ p = .013,\ d = -.5136$, (95% CI [-0.91, -0.10]). However, robot evaluation was only marginally higher in the social

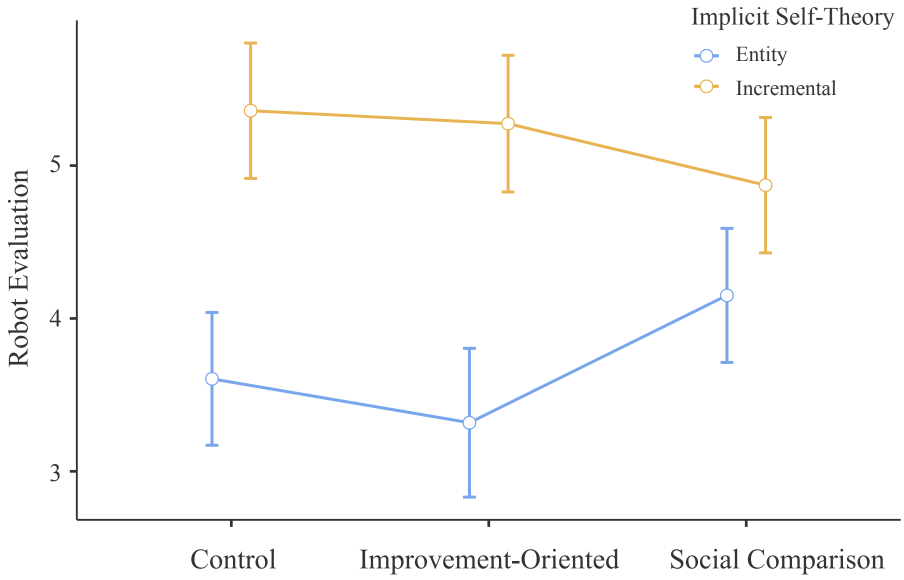


Figure 4.5: Robot Evaluation means and 95% confidence intervals for the effects of implicit self-theories and feedback.

comparison condition than in the control condition ($M = 3.6$, $SD = 1.83$), $t = -1.741$, $p = .083$, $d = -0.3366$, (95% CI [-0.71, 0.04]).

Planned comparison results for incremental theorists, revealed that improvement-oriented feedback ($M = 5.27$, $SD = 1.53$) led to a nonsignificantly more favorable robot evaluation than social comparison feedback ($M = 4.87$, $SD = 1.73$), $t(299) = 1.26$, $p = .209$, $d = 0.2483$, (95% CI [-0.14, 0.63]). Furthermore, no significant difference in robot evaluation was found between the improvement-oriented condition and the control condition ($M = 5.50$, $SD = 1.15$), $t(299) = 0.264$, $p = .792$, $d = 0.0521$, (95% CI [0.33, 0.43]). Given these results, H2 was partially supported, and H3 was not supported (see Figure. 4.5).

Identity Threat (H4)

A 2 (implicit self-theory) ×3 (feedback) between-subjects ANOVA on identity threat was also conducted. As expected, a significant main effect was observed for implicit self-theory ($F(1, 299) = 15.52$, $p < .001$). There were no predictions made regarding the effect of feedback condition, nor was there an interaction predicted. No significant effect of feedback condition, ($F(2, 299) = 0.09$, $p = .918$), and no significant interaction emerged ($F(2, 299) = 1.15$, $p = .318$). Planned comparisons revealed that incremental theorists perceived robots to be significantly less threatening to human identity and uniqueness ($M = 3.78$, $SD = 1.71$) than entity theorists ($M = 4.53$, $SD = 1.58$), $t(299) = 3.94$, $p < .001$, $d = 0.452$, (95% CI [0.22, 0.68]). This result supports H4 (see Figure. 4.6).

Realistic Threat (H5)

A similar ANOVA test was conducted on realistic threat. As expected, a significant effect for implicit self-theory, on realistic threat ($F(1, 299) = 17.428$, $p < .001$), was noted. There were no predictions made regarding the effect of feedback condition,

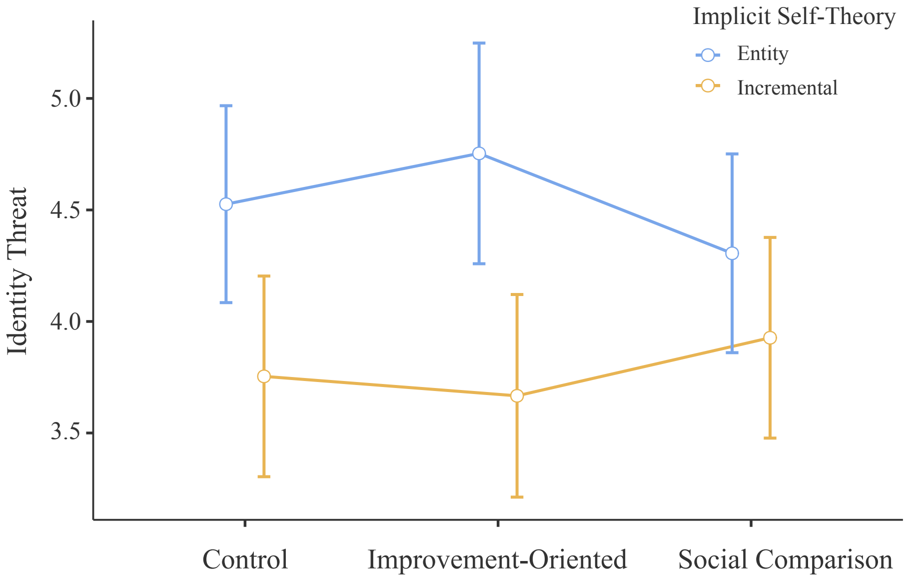


Figure 4.6: Identity Threat means and 95% confidence intervals for the effects of implicit self-theories and feedback.

nor was there an interaction predicted. No significant effect of feedback condition ($F(2, 299) = 0.61$, $p = .542$), was observed, however an unpredicted interaction between implicit self-theory and feedback emerged ($F(2, 299) - = 4.74$, $p < .009$). Planned comparisons showed that entity theorists perceived robots to pose significantly more threat to employment, resources, and well-being ($M = 4.78$, $SD = 1.14$) than incremental theorists ($M = 4.19$, $SD = 1.35$), $t(299) = 4.17$, $p < .001$, $d = 0.479$, (95% CI [0.25, 0.70]). This finding supports H5 (see Figure. 4.7).

4.5 Discussion

The results of Study 3 indicate that incremental theorists (with their propensity for developing ability and redressing failure) are more likely to indicate an interest in playing against an intelligent robot after imagining losing to it, thus supporting Hypothesis 1. Curiously however, entity theorists (who are thought to experience positive affect after comparing themselves to worse performers following failure), did not evaluate the robot significantly more favorably in the social comparison condition than in the control condition. As well, incremental theorists were not influenced by feedback, in so much as they evaluated the robot favorably regardless of the feedback they received. This finding, in particular, does not appear to be entirely consistent with the broader implicit self-theory literature, linking improvement-oriented feedback to incremental theories (Cimpian et al., 2007; Haimovitz and Henderlong Corpus, 2011; Kamins and Dweck, 1999; Pomerantz and Kempner, 2013; Skipper and Douglas, 2012) and the tendency to make positive judgments (i.e., Lou and Noels, 2020), in consequence. However, this finding is consistent with that of Study 2 which showed that incremental theorists are unaffected by robot-delivered praise and tend to respond more favorably to social robots in general. Consequently, Hypothesis 2 and Hypothesis 3 were not completely supported. That said, both findings were generally in the predicted direction, which might imply that a larger sample would

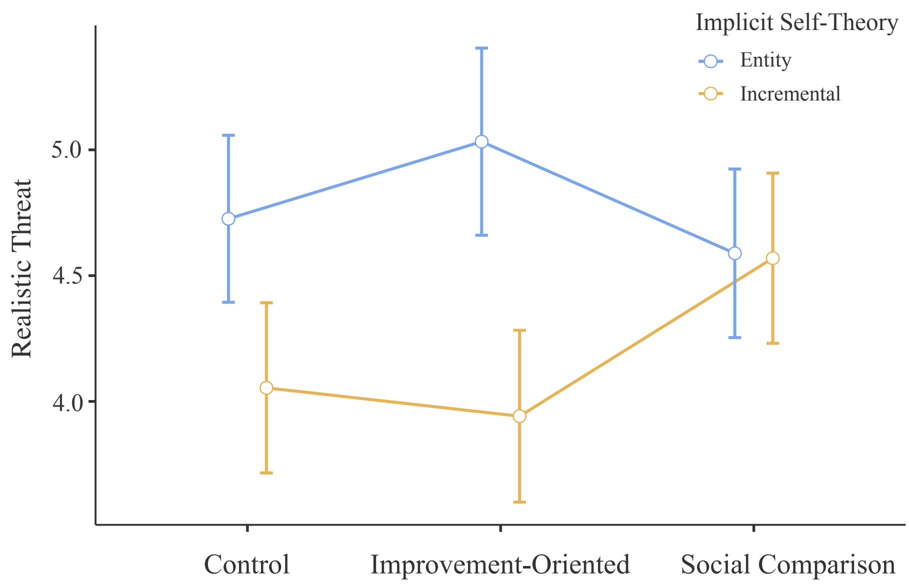


Figure 4.7: Realistic Threat means and 95% confidence intervals for the effects of implicit self-theories and feedback.

potentially confirm these effects. Consistent with Hypothesis 4, however, the results showed that compared to entity theorists, incremental theorists rated the robot as posing less identity relevant threats.

It was also observed that entity theorists relative to incremental theorists rated the robot as presenting more realistic threats, in support of Hypothesis 5.

Unexpectedly, an unpredicted interaction was suggested between incremental theory and social comparison feedback on realistic threat. Plausible reasons for this interaction are not evident, thereby necessitating further research to investigate this finding and to tease out possible explanations.

Moreover, Study 3 addressed the two shortcomings noted in Study 2, Section 3.5.1. One is the use of a Baxter Robot (vs. a NAO V5 robot), which indicates that the impact of individual implicit self-theory on one's responses to social robots is apparent in regard to at least one other robot type. The second, is the use of a control condition as a baseline.

4.5.1 Limitations and Future Directions

Generalizations from these findings are of course limited by the robot that was used. In other words, a Baxter Robot, which as a six foot three inch humanoid robot, cuts an imposing figure (see Figure. 4.1). The obvious question then is whether the results reported in this chapter would be observed with different robot types. Future work using alternative aesthetic robot forms (e.g., mechanomorphic, android), can shed light on this question.

Crucially, more research is needed to determine whether the effects noted here emerge following a real-life human-robot interaction. Although recent work has found fairly strong correlations of findings between online and in-person experiments (Babel et al., 2021), the imagining of robot delivered feedback following a perceived loss, is hardly the same thing as receiving direct feedback from a robot (Hoffman and Zhao, 2020) in a real-life encounter (especially after competing against and losing to it). One should bear in mind however, that this study was conducted before and during a COVID-19 lockdown, which severely impeded any possibility of running a physical follow-up experiment at the time. Nevertheless, as physical experiments become more feasible again, it is important that prospective research uses an in-person experimental design.

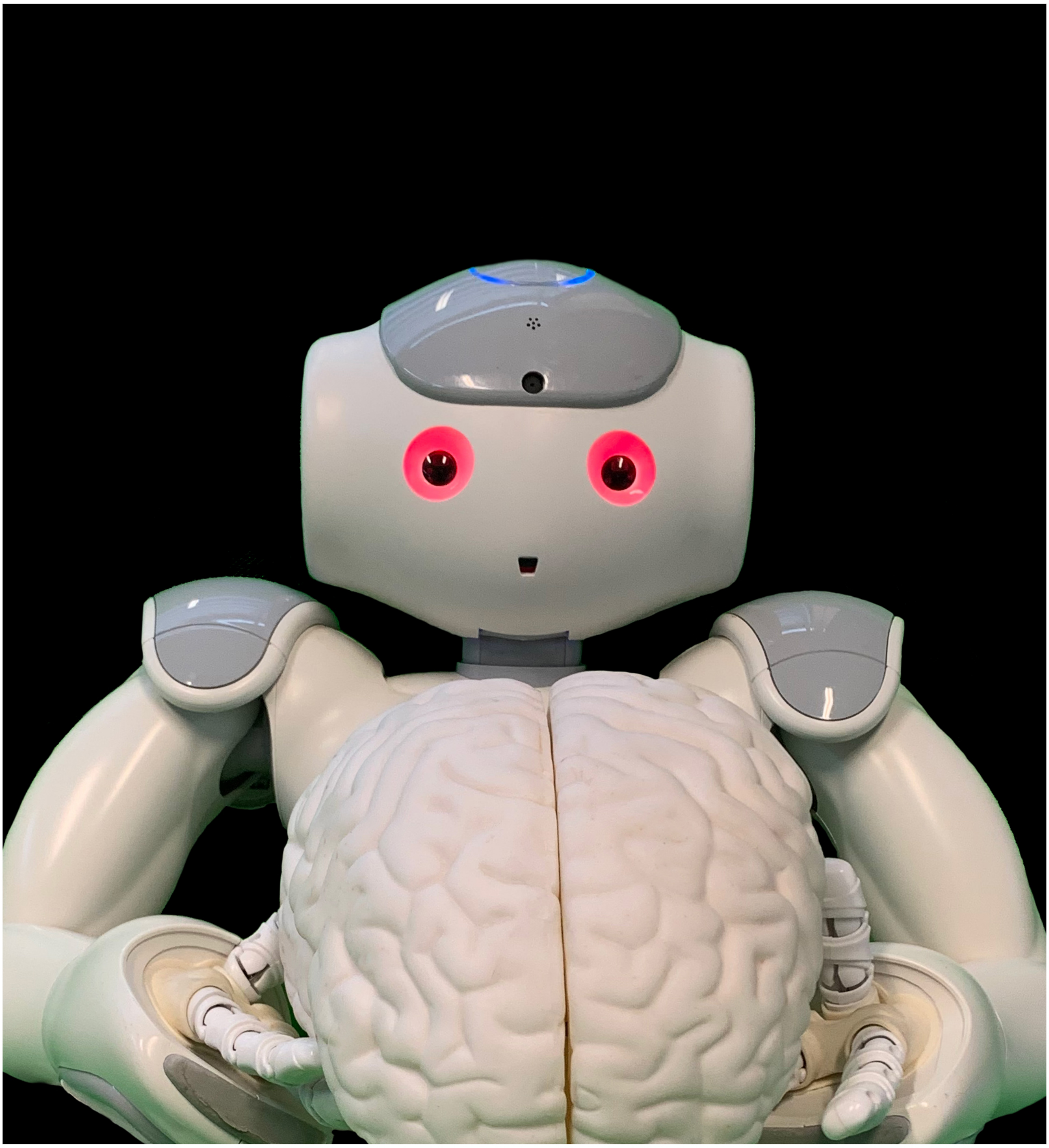
Future studies might also include a human control condition comparable to the robot condition in terms of general knowledge and perceived intelligence (e.g., a professional quizzer).

4.6 Conclusion

Overall, the findings detailed in this chapter offer further support for the proposition that implicit self-theory serves as an important variable in determining how individuals respond to robots. As expected, incremental theorists were rather receptive to robots that could both physically and mentally outperform human beings. In contrast, entity theorists viewed such robots as threatening, and subsequently rated them unfavorably. As such, these data demonstrate the first research evidence that implicit self-theory is a determinant of whether one person experiences threat perception (i.e., realistic and identity threats) to a greater degree than another. The next chapter summarizes the findings of the studies reported in this thesis. It outlines limitations and suggests possible future directions for research. It concludes with closing remarks.

Chapter V

General Discussion



V

General Discussion

5.1 Introduction

This chapter concludes the thesis and outlines suggestions for further research. Correspondingly, this chapter is organised as follows. Section 5.2 summarizes the findings reported herein, while Section 5.3 presents the theoretical contributions and potential applications. Section 5.4 discusses limitations and directions for future work. Finally, a conclusion is given in Section 5.5.

5.2 Summary of Findings

Three studies demonstrate that implicit self-theories are an important variable that influences the manner in which people perceive, evaluate, and respond to, social robots. Study 1 presented initial data suggesting that implicit self-theories do indeed exert some influence over individuals' perceptions of social robots. More precisely, compared to incremental theorists, entity theorists expressed greater robot anxiety. This result held even when controlling for well-known covariate influences, like prior robot experience, media exposure to science fiction, technology commitment, and certain demographic factors. That said, only marginal effects were obtained for both attitudinal and intentional robot acceptance, respectively. In addition, it was found that incremental theorists responded more favorably to social robots, compared to entity theorists. Finally, and interestingly, the results of Study 1 found that entity theorists reported more favorable responses to a social robot positioned as a servant (vs. an assistant).

Study 2 built on the findings of Study 1, and examined the interactive influence of individual implicit self-theories, and robot-delivered praise, on people's evaluation of a social robot after success on a difficult task. Results showed that, those embracing more of an entity theory evaluated a robot lower in both perceived intelligence and likability when it praised them for effort (vs. ability). Additionally, Study 2 revealed that those embracing more of an incremental theory remained relatively unaffected by either praise type, and instead evaluated the robot higher in both perceived intelligence and likability irrespective of the praise it delivered.

Study 3 extended and complemented the findings of Studies 1 and 2 by investigating the effects of implicit self-theories on people's responses to an intelligent robot after experiencing a failure of sorts. Specifically, participants watched a video of an apparently autonomous intelligent robot defeating human quiz players in a general knowledge game. Following the video, participants received either social-comparison feedback, improvement-oriented feedback, or no feedback, and were subsequently given the opportunity to play against the robot. The results of Study 3 indicated that those endorsing an incremental theory were more likely to indicate an interest in playing against an intelligent robot after imagining losing to it, as well as to report less identity threats (i.e., threats to human identity and distinctiveness) than those endorsing more of an entity theory. Against predictions (yet consistent with the findings of Study 2), incremental theorists were found to express highly favorable responses to an intelligent autonomous robot, regardless of the feedback they were presented with. Equally unexpected is the finding that, entity theorists did not rate the robot significantly differently in the social comparison feedback condition than in the control feedback condition. Strikingly, results showed that entity theorists relative to incremental theorists perceived autonomous intelligent robots to be significantly more threatening, both in terms of identity threats and realistic threats (i.e., threats to safety, jobs, and resources).

It is appropriate to reiterate here that these studies assessed implicit self-theories with: (1) a variety of issues relevant to HRI (i.e., public perception of emerging social robots and robot acceptance, how robot-delivered praise impacts evaluations of likability and perceived intelligence, and how people respond to an intelligent robot that outperforms human beings), (2) across different self-theory conducing interaction contexts (i.e., success/failure, challenges/setbacks), (3) scenario media (i.e., text, in-person, and video), (4) using different robot types (i.e., a NAO robot and a Baxter Robot), (5) an array of established HRI measures (e.g., NARS, perceived intelligence, likability, identity threat, realistic threat), and finally (6) different methods for assessing implicit self-theories (i.e., both measured and manipulated).

Collectively, the above might form a basis from which one can draw reasonably substantive (though early) conclusions about the role of implicit self-theories in HRI.

That said, although these findings are largely consistent with the hypotheses, there were some exceptions as noted above.[1] However, as other researchers have noted (e.g., Smith et al., 2021), this is to be expected, at least to some extent, when expanding social psychological theories beyond their intended scope and into the domain of HRI (see also Wullenkord and Eyssel, 2020).

Considered together however, these findings expand the state-of-the-art, by providing the first empirical support, for the proposition that implicit self-theory influences, people's perceptions, evaluations, and responses toward social robots. As such, these findings offer novel theoretical and practical contributions which will be discussed next.

5.3 Theoretical Contributions and Applications

This research makes several original contributions to multiple literatures. First, this research contributes the literature on robot anxiety in HRI (Nomura et al., 2004; Syrdal et al., 2009; de Graaf and Allouch, 2013; Kanero et al., 2021) by establishing

[1] As has been described in previous chapters, all such findings followed the expected pattern and were either marginally significant or barely missed statistical significance, thereby suggesting that a larger sample might well confirm the predicted effects.

that entity theorists report greater robot anxiety than incremental theorists (Study 1). Likewise, this research enriches the HRI literature with respect to work on robot acceptance (e.g., De Graaf et al., 2015; Ezer, 2008; Bernotat and Eyssel, 2018; de Graaf et al., 2019; Harrington et al., 2021; Troncone et al., 2020; de Kervenoael et al., 2020; Ghazali et al., 2020; Whelan and Casey, 2021) by demonstrating that incremental (vs. entity) theorists exhibit more favorable responses to social robots overall (Study 1), regardless of the type of praise a robot delivers (Study 2), or whether it can physically and mentally outperform human beings (Study 3).

Results also contribute to the literature on robot-delivered feedback (e.g., Akalin et al., 2019; Fasola and Matarić, 2013; Ghazali et al., 2019; Schermerhorn et al., 2008; Leite et al., 2012; Swift-Spong et al., 2015; Leyzberg et al., 2014; Ham and Midden, 2014) by showing that one's response to robot-delivered praise (ability, effort-focused) is influenced by one's implicit self-theory orientation (Study 2). More specifically, in identifying the significant differential effects of robot praise on two core dimensions of robot acceptance: likability and perceived intelligence for entity theorists.

Furthermore, these findings (principally, the results of Study 3) expand the literature on perceived threat in HRI. As described in Chapter 4, this line of research finds that autonomous and intelligent robots, and more generally, "smart" machines, incite threats to human uniqueness, safety, and resources, which in turn, impedes positive perceptions of, and willingness to engage with, such technology (Ferrari et al., 2016; Yogeeswaran et al., 2016; Złotowski et al., 2017; Meissner et al., 2020; Kieslich et al., 2021; Huang et al., 2021; Hover et al., 2021). Some, such as Złotowski and colleagues (2017), have posited that such threats, "may, be universal and experienced by all people since it concerns humankind uniqueness and distinctiveness as a whole" (p.53). The findings of Study 3 suggest that this is not necessarily the case. Rather, these findings indicate that significant individual variability exists, such that incremental theorists are rather receptive to robots that can both physically and mentally outperform human beings. In contrast, entity theorists view such robots as threatening and hence rate them unfavorably.

This work also broadens the social psychological literature on implicit self-theories beyond domains such as education and consumer behavior (reviewed in Chapter 1), into the domain of HRI and social robotics. The results of Study 2 further extend this body of work by presenting what might be the first evidence of a strong and direct interactive match between implicit self-theory (entity vs. incremental) and praise (ability vs. effort), outside of an educational or learning-related context.

Over and above their theoretical contributions, these findings may hold practical value for HRI designers interested in developing social robotic products that are more humanly engaging. For example, the findings of Study 2 seem to indicate that a robot specially designed to praise for ability (e.g., "You're great at this" or "You're so smart"), perhaps from time to time sporadically—could provide considerable enjoyment, satisfaction, and engagement for entity theorists, and at the same time not diminish the user experience for incremental theorists.

Subsequently, the results of Study 3 would suggest that a robot that is presented as intelligent and capable of outperforming human beings could achieve greater acceptance and use among incremental theorists (vs. entity theorists). Whereas a robot that is presented as being less intelligent and unlikely to outperform human beings could increase the probability of acceptance among entity theorists, for example.

As well, the combined results of Studies 2 and 3 illustrate how the relative salience of a particular self-theory, and by extension, one's experience of social robots, can be activated purposively by exposing participants to text that emphasize either the malleability (for an incremental theorist) or fixedness (for an entity theorist) of self-attributes. Resultantly, these findings may offer direction to marketers of social robotic products. Recall, the findings of Study 1 demonstrate that only entity and not incremental theorists are likely to experience strong aversion toward social robotic products designed for everyday use. Accordingly, marketers may consider influencing potential consumers' implicit self-theories toward more of an incremental view through their marketing stimuli, which may elicit greater feelings of acceptance and the likelihood of assent.

5.4 Limitations and Future Directions

The contributions of the present studies are qualified by certain limitations (beyond those already described in previous chapters).

One key limitation (although not within the scope of the planned hypotheses), might be that all well-known covariate influences common to HRI research were not tested, e.g., individual differences in age (Kuo et al., 2009), culture (Bartneck et al., 2007; Bernotat and Eyssel, 2018; Haring et al., 2014), mind perception (Gray and Wegner, 2012), personality traits (Santamaria and Nathan-Roberts, 2017), anthropomorphism (Duffy, 2003), and occupation (Gnambs and Appel, 2019). It would be useful then for future studies to examine how such variables may relate to, or interact with implicit self-theories.

Another potential limitation might be that every participant who was recruited into Studies 2 and 3 was required, in part, to have no experience with robotics. It is worth noting, however, that a recent review by Naneva and colleagues (2020) suggests, that many people have yet to have contact with social robots (let alone experience with robotics, more broadly). Thus, as the authors seem to conclude, studies excluding these individuals may be more generalizable than initially thought.

A related limitation is that all participants resided in an English-speaking country (Studies 1-3), and might therefore be classified as "WEIRD" (Western, educated, industrialized, rich, and democratic; see Henrich et al., 2010; Apicella et al., 2020; Rauthmann, 2020; Linxen et al., 2021). Consequently, these exclusion criteria limit the external validity of findings. Therefore, future studies might benefit from using different eligibility criteria.

Moreover, the findings of Studies 1 and 3 are obscured by the samples used, which were drawn from Mturk. Although extant research suggests Mturk samples produce valid and reliable data (Bartneck et al., 2015; Buhrmester et al., 2016; Huff and Tingley, 2015) of equal or greater quality than student subject pools (Kees et al., 2017), there is a small but growing body of research that has documented many troubling trends regarding samples obtained from MTurk (see Walter et al., 2019; Peer et al., 2021). These include, but are not limited to, participant carelessness (Pyo and Maxfield, 2021), insufficient effort (Lu et al., 2021), and fraudulent responding (Agley et al., 2021). An especially worrisome finding that has emerged in recent years, is that MTurk samples may respond differently depending on the day of the week (see Binder, 2022; Arechar et al., 2017) and time of the day (see Fordsham et al., 2019; Casey et al., 2017) of participation. While these studies are few in number, and more research is clearly needed, they nonetheless point to potential threats (i.e., generalizability, reliability, and even replication of findings) to MTurk data integrity (e.g., Casey et al., 2017). Hence, although measures were taken to strengthen the quality of these data, such as placing attention and validity checks in the survey (e.g., Robinson et al., 2019; Cobanoglu et al., 2021), it is still a possibility that some respondents may have provided dishonest, imprecise, or otherwise poor quality responses (see Barends and de Vries, 2019; Dennis et al., 2020).

On top of this, Mturk samples are shown to be more technologically-savvy (see Munger et al., 2018), and more well-educated (see Aruguete et al., 2019) than the general public, which was confirmed in these data. Specifically, participants in Study 1 were relatively astute with respect to general technology use, which has been previously associated with favorable attitudes toward social robots (Reich-Stiebert and Eyssel, 2015).

Relatedly, all samples consisted of highly educated individuals, with the vast majority in any given sample having some college, or greater, education. Previous research has shown that more educated individuals relative to less educated individuals, tend to hold more positive views about robots (Gnambs and Appel, 2019).

Another issue worth raising is that every participant self-selected to take part in research about social robots. Thus, it may be the case that each sample was subject to self-selection bias. In consequence, their view of social robots could differ from those of the general population.

An additional limitation across all studies, is that the measures used were obtained via self-report, which although theoretically substantiated and practically applicable, may be vulnerable to bias (e.g., participants may provide self-enhancing responses; Podsakoff et al., 2003). Therefore, future research employing alternative samples, and more objective measures (e.g., use of an electroencephalogram; Bossi et al., 2020) seems both warranted and potentially beneficial.

Another limitation of the present research might be the use of scenarios that were designed to manifest most markedly, implicit self-theory-led experiences (i.e., those pertaining to challenge, success, failure, and setbacks). Future research could extend this investigation by employing scenarios less conducive to stimulating implicit self-theories. Quite possibly, usability studies, grounded in engineering (e.g., those validating robotic algorithms), may be suitable for this purpose.

One final limitation relates to the numerous difficulties in performing physical experiments during the COVID-19 pandemic (see Feil-Seifer et al., 2020; Mirsky et al., 2021). Indeed, it should be acknowledged that the option of running in-person experiments was severely restricted, and in multiple cases impossible (e.g., sudden and repeated national lockdowns, physical distancing, face covering mandates, and stay-at-home orders) at the time at which this research was conducted.

Despite these limitations, the present findings open up avenues for future work. One interesting avenue for researchers to explore might be, the relationship between one's implicit self-theory orientation and their perceptions of trust toward social robots (Schaefer, 2016; Kok and Soh, 2020; Lewis et al., 2018; Schüle et al., 2022).

Past research posits, for example, that an individuals' trust in a robot is diluted by any functional, mechanical, or programming errors it might exhibit (e.g., Rossi et al., 2017, 2018; Robinette et al., 2017). In such instances, it is entirely possible that future research may find that entity theorists' who focus on outcome (Levy and

Dweck, 1998), and are less tolerant of transgressions (Haselhuhn et al., 2010), will lose trust in a robot that displays such errors (Park and John, 2018). Comparatively, incremental theorists who focus on learning goals (Molden and Dweck, 2006), and value effort processes (Levy and Dweck, 1998) over flawless performance (Dweck, 2017), will be less likely to lose trust in a robot that errs. Future studies could test these predictions by having individuals interact with a social robot that makes errors (e.g., Rossi et al., 2020), and by assessing implicit self-theory and trust evaluations (e.g., Haselhuhn et al., 2017).

Another intriguing area to be explored might be whether and how implicit self-theories (incremental vs. entity), and robot-delivered negative feedback, may interactively increase performance, and the extent to which people assign more (vs. less) favorable robot evaluations and use-intentions, in response. To illustrate: There is some evidence that a robot's negative or impolite feedback can engender behavior change (see Midden and Ham, 2009; Ham and Midden, 2014) and increased task performance (Rea et al., 2021), although people tend to prefer a robot's positive feedback (Ghazali et al., 2019). Given that incremental theorists, compared to entity theorists, exhibit positive effort beliefs in the face of adversity (Hong et al., 1999), and are more motivated by learning goals (Dweck and Leggett, 1988), while favoring goal progress cues (Mathur et al., 2014), it could be that incremental (vs. entity) theorists respond somewhat positively to a robot that displays negative or critical feedback to encourage better performance. In any case, future researchers could test this prediction by implementing a human-robot interaction in the context of a robot-assisted training task (e.g., Rea et al., 2021), and by measuring implicit self-theory and different measures of robot acceptance, as in Study 1.

Future work should also examine the consequences of implicit self-theories in the context of everyday human-robot interactions (i.e., field studies). For example, researchers could investigate whether—and to what extent—individual implicit self-theories bear on long-term human-robot cohabitation (e.g., Dziergwa et al., 2018) and longitudinal interactions (Hart et al., 2022; Irfan et al., 2022; Zhao and McEwen, 2022; Ostrowski et al., 2022).

Another interesting area to be explored is whether implicit self-theories promote or reduce individuals' proclivity to exhibit aggressive behaviors towards robots (Bartneck and Hu, 2008; Nomura et al., 2016). The findings of Study 3 imply that an entity theory precipitates greater identity and realistic threats in response to autonomous intelligent robots. As demonstrated in Study 1, entity beliefs are associated with greater negative attitudes regarding robots, earlier research has linked entity beliefs with aggressive reactions to ambiguous provocations (Yeager et al., 2013). It seems reasonable to propose then, that entity theorists may be prone to acts of robot-directed aggression (Salvini et al., 2010). Contrastingly, incremental theorists who tend to respond more positively to robots (Studies 1-3) may be more likely to intervene when a robot is being mistreated or abused (Tan et al., 2018). Of course, future research is needed to empirically test these propositions.

5.5 Conclusion

Taken together, the pattern of findings converges on the conclusion that implicit self-theories are indeed an important variable that influences the manner in which people perceive, evaluate, and respond to, social robots. As such, this research may be considered a timely and relevant contribution, given the expected rise of robots in the social domain of everyday experience (e.g., Haegele, 2016; Beraldo et al., 2019; Galin and Mamchenko, 2021; de la Feria and Grau Ruiz, 2022), in combination with the increasing demand for research on the psychological variables and individual difference factors underlying HRI (e.g., Eyssel, 2017; Matthews et al., 2020; Collins, 2019; Morsunbul, 2019; Xu, 2019; Robert et al., 2020; Brylska et al., 2022). It is thus a feasible (or, perhaps more accurately, a hopeful) possibility that as the field moves forward in an ongoing quest for defining and informing the design of highly effective social robots, that these findings (and future research based on these findings) may find application in entirely new kinds of social robots. That is, robots capable of reflecting or shaping self-beliefs, and consequently, robots that human beings willingly adopt in their lives, as extensions of their identity. At the very least this research points to implicit self-theories as an important and meaningful variable with which to further advance HRI research and praxis.

Appendix A

Copyright Acknowledgements

Parts of this work have been published under the following titles:

Allan, D. D., Vonasch, A. J., and Bartneck, C. (2022b). The doors of social robot perception: The influence of implicit self-theories. *International Journal of Social Robotics*, 14(1):127–140

Allan, D. D., Vonasch, A. J., and Bartneck, C. (2022c). "i have to praise you like i should?" the effects of implicit self-theories and robot-delivered praise on evaluations of a social robot. *International Journal of Social Robotics*

Allan, D. D., Vonasch, A. J., and Bartneck, C. (2022a). Better than us: The role of implicit self-theories in determining perceived threat responses in hri. In *Proceedings of the 2022 ACM/IEEE International Conference on Human-Robot Interaction*, pages 215–224

Appendix B

Data Availability

All datasets, including supplementary material, are available via the Open Science Framework and can be accessed at the following links:

Chapter 2: https://osf.io/m84q7/?view_only=db98cf9f5e334ca7aca9349799ef12df

Chapter 3: https://osf.io/v5gek/?view_only=0621e70677fe4362b796a0afce165981

Chapter 4: https://osf.io/xefj4/?view_only=1499ac3f73424c2481634c3212a6de66

Appendix C

Ethics Declarations

This work complies with the ethical standards of the University Canterbury. Informed consent was obtained from every participant involved in the studies reported herein, including, where appropriate, explicit written permission to use their image and likeness. The research protocol for Studies 1-3 were reviewed and approved by the Human Research Ethics Committee of the University of Canterbury (HREC) and may be verified using the following references:

Study 1: (HEC 2019/53/LR-PS)

Study 2: (HEC 2020/130)

Study 3: (HEC 2021/33)

References

Ablard, K. E. (2002). Achievement goals and implicit theories of intelligence among academically talented students. *Journal for the Education of the Gifted*, 25(3):215–232.

Ablard, K. E. and Mills, C. J. (1996). Implicit theories of intelligence and self-perceptions of academically talented adolescents and children. *Journal of Youth and Adolescence*, 25(2):137–148.

Agley, J., Xiao, Y., Nolan, R., and Golzarri-Arroyo, L. (2021). Quality control questions on amazon's mechanical turk (mturk): A randomized trial of impact on the usaudit, phq-9, and gad-7. *Behavior Research Methods*.

Aiken, L. S., West, S. G., and Reno, R. R. (1991). *Multiple regression: Testing and interpreting interactions*. sage.

Akalin, N., Kristoffersson, A., and Loutfi, A. (2019). The influence of feedback type in robot-assisted training. *Multimodal Technologies and Interaction*, 3(4):67.

Aldahdouh, T. Z., Nokelainen, P., and Korhonen, V. (2018). Innovativeness of staff in higher education:" do implicit theories and goal orientations matter?". *International Journal of Higher Education*, 7(2):43–57.

Ali, S., Devasia, N., and Breazeal, C. (2021). Designing games for enabling co-creation with social agents. *arXiv preprint arXiv:2105.00201*.

Aljukhadar, M. and Senecal, S. (2021). The effect of consumer-activated mind-set and product involvement on the compliance with recommender system advice. *SAGE Open*, 11(3):21582440211031550.

Allan, D. D., Vonasch, A. J., and Bartneck, C. (2022a). Better than us: The role of implicit self-theories in determining perceived threat responses in hri. In *Proceedings of the 2022 ACM/IEEE International Conference on Human-Robot Interaction*, pages 215–224.

Allan, D. D., Vonasch, A. J., and Bartneck, C. (2022b). The doors of social robot perception: The influence of implicit self-theories. *International Journal of Social Robotics*, 14(1):127–140.

Allan, D. D., Vonasch, A. J., and Bartneck, C. (2022c). "i have to praise you like i should?" the effects of implicit self-theories and robot-delivered praise on evaluations of a social robot. *International Journal of Social Robotics*.

Andrist, S., Mutlu, B., and Tapus, A. (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*, CHI '15, pages 3603–3612, New York, NY, USA. ACM.

Apicella, C., Norenzayan, A., and Henrich, J. (2020). Beyond weird: A review of the last decade and a look ahead to the global laboratory of the future. *Evolution and Human Behavior*, 41(5):319–329.

Appel, M., Izydorczyk, D., Weber, S., Mara, M., and Lischetzke, T. (2020). The uncanny of mind in a machine: Humanoid robots as tools, agents, and experiencers. *Computers in Human Behavior*, 102:274–286.

Arechar, A. A., Kraft-Todd, G. T., and Rand, D. G. (2017). Turking overtime: how participant characteristics and behavior vary over time and day on amazon mechanical turk. *Journal of the Economic Science Association*, 3(1):1–11.

Aronson, J., Fried, C. B., and Good, C. (2002). Reducing the effects of stereotype threat on african american college students by shaping theories of intelligence. *Journal of Experimental Social Psychology*, 38(2):113–125.

Aruguete, M. S., Huynh, H., Browne, B. L., Jurs, B., Flint, E., and McCutcheon, L. E. (2019). How serious is the 'carelessness' problem on mechanical turk? *International Journal of Social Research Methodology*, 22(5):441–449.

Babel, F., Kraus, J., Hock, P., Asenbauer, H., and Baumann, M. (2021). Investigating the validity of online robot evaluations: Comparison of findings from an one-sample online and laboratory study. In *Companion of the 2021 ACM/IEEE International Conference on Human-Robot Interaction*, page 116–120.

Barends, A. J. and de Vries, R. E. (2019). Noncompliant responding: Comparing exclusion criteria in mturk personality research to improve data quality. *Personality and individual differences*, 143:84–89.

Bartneck, C., Belpaeme, T., Eyssel, F., Kanda, T., Keijsers, M., and Šabanović, S. (2020). *Human-Robot Interaction: An Introduction.* Cambridge University Press, Cambridge.

Bartneck, C., Croft, E., and Kulic, D. (2009). Measurement instruments for the anthropomorphism, animacy, likeability, perceived intelligence, and perceived safety of robots. *International Journal of Social Robotics*, 1(1):71–81.

Bartneck, C., Duenser, A., Moltchanova, E., and Zawieska, K. (2015). Comparing the similarity of responses received from studies in amazon's mechanical turk to studies conducted online and with direct recruitment. *PLoS One*, 10(4):e0121595.

Bartneck, C. and Forlizzi, J. (2004). A design-centred framework for social human-robot interaction. In *RO-MAN 2004. 13th IEEE International Workshop on Robot and Human Interactive Communication (IEEE Catalog No. 04TH8759)*, pages 591–594. IEEE.

Bartneck, C. and Hu, J. (2008). Exploring the abuse of robots. *Interaction Studies*, 9(3):415–433.

Bartneck, C. and Keijsers, M. (2020). The morality of abusing a robot. *Paladyn, Journal of Behavioral Robotics*, 11(1):271–283.

Bartneck, C., Reichenbach, J., and Carpenter, J. (2006). Use of praise and punishment in human-robot collaborative teams. In *ROMAN 2006-The 15th IEEE International Symposium on Robot and Human Interactive Communication*, pages 177–182. IEEE.

Bartneck, C., Reichenbach, J., and Carpenter, J. (2008). The carrot and the stick: The role of praise and punishment in human–robot interaction. *Interaction Studies*, 9(2):179–203.

Bartneck, C., Suzuki, T., Kanda, T., and Nomura, T. (2007). The influence of people's culture and prior experiences with aibo on their attitude towards robots. *Ai & Society*, 21(1-2):217–230.

Belpaeme, T., Young, J., Gunes, H., and Riek, L., editors (2020). *HRI '20: Proceedings of the 2020 ACM/IEEE International Conference on Human-Robot Interaction*, New York, NY, USA. Association for Computing Machinery.

Beraldo, G., Di Battista, S., Badaloni, S., Menegatti, E., and Pivetti, M. (2019). Sex differences in expectations and perception of a social robot. In *2018 IEEE Workshop on Advanced Robotics and its Social Impacts (ARSO)*, pages 38–43. IEEE.

Bergen, R. S. (1991). *Beliefs about intelligence and achievement-related behaviors.* Unpublished doctoral thesis, University of Illinois at Urbana-Champaign.

Bernotat, J. and Eyssel, F. (2018). Can ('t) wait to have a robot at home?-japanese and german users' attitudes toward service robots in smart homes. In *2018 27th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*, pages 15–22. IEEE.

Billard, A. and Dautenhahn, K. (1997). Grounding communication in situated, social robots. In *Proceedings Towards Intelligent Mobile Robots Conference, Report No. UMCS-97-9-1, Department of Computer Science, Manchester University.*

Binder, C. C. (2022). Time-of-day and day-of-week variations in amazon mechanical turk survey responses. *Journal of Macroeconomics*, 71:103378.

Blackwell, L. S., Trzesniewski, K. H., and Dweck, C. S. (2007). Implicit theories of intelligence predict achievement across an adolescent transition: A longitudinal study and an intervention. *Child development*, 78(1):246–263.

Bossi, F., Willemse, C., Cavazza, J., Marchesi, S., Murino, V., and Wykowska, A. (2020). The human brain reveals resting state activity patterns that are predictive of biases in attitudes toward robots. *Science robotics*, 5(46).

Bostwick, K. C. and Becker-Blease, K. A. (2018). Quick, easy mindset intervention can boost academic achievement in large introductory psychology classes. *Psychology Learning & Teaching*, 17(2):177–193.

Breazeal, C. (2003). Emotion and sociable humanoid robots. *International journal of human-computer studies*, 59(1-2):119–155.

Brondi, S., Pivetti, M., Di Battista, S., and Sarrica, M. (2021a). What do we expect from robots? social representations, attitudes and evaluations of robots in daily life. *Technology in Society*, 66:101663.

Brondi, S., Pivetti, M., Di Battista, S., and Sarrica, M. (2021b). What do we expect from robots? social representations, attitudes and evaluations of robots in daily life. *Technology in Society*, 66:101663.

Brummelman, E. (2020). *Psychological Perspectives on Praise*. Routledge.

Brylska, P., Cobanoglu, C., and Dogan, S. (2022). Users versus non-users: The impact of experience on hotel guests' attitudes towards service robots in hotels. In *Information and Communication Technologies in Tourism 2022*, pages 55–66. Springer International Publishing.

Buhrmester, M., Kwang, T., and Gosling, S. D. (2016). *Amazon's Mechanical Turk: A new source of inexpensive, yet high-quality data?* Methodological issues and strategies in clinical research, 4th ed. American Psychological Association, Washington, DC, US.

Burgoyne, A. P., Hambrick, D. Z., and Macnamara, B. N. (2020). How firm are the foundations of mind-set theory? the claims appear stronger than the evidence. *Psychological Science*, 31(3):258–267.

Burnette, J. L., O'Boyle, E. H., VanEpps, E. M., Pollack, J. M., and Finkel, E. J. (2013). Mind-sets matter: A meta-analytic review of implicit theories and self-regulation. *Psychological bulletin*, 139(3):655.

Cagiltay, B., Ho, H.-R., Michaelis, J. E., and Mutlu, B. (2020). Investigating family perceptions and design preferences for an in-home robot. In *Proceedings of the Interaction Design and Children Conference*, pages 229–242.

Carpinella, C. M., Wyman, A. B., Perez, M. A., and Stroessner, S. J. (2017). The robotic social attributes scale (rosas) development and validation. In *Proceedings of the 2017 ACM/IEEE International Conference on human-robot interaction*, pages 254–262.

Casey, L. S., Chandler, J., Levine, A. S., Proctor, A., and Strolovitch, D. Z. (2017). Intertemporal differences among mturk workers: Time-based sample variations and implications for online data collection. *SAGE Open*, 7(2):2158244017712774.

Cha, Y., Baek, S., Ahn, G., Lee, H., Lee, B., Shin, J.-e., and Jang, D. (2020). Compensating for the loss of human distinctiveness: The use of social creativity under human–machine comparisons. *Computers in Human Behavior*, 103:80–90.

Chandler, J., Paolacci, G., and Mueller, P. (2013). *Risks and rewards of crowdsourcing marketplaces*, pages 377–392. Springer.

Chandrasekar, R. (2014). Elementary? question answering, ibm's watson, and the jeopardy! challenge. *Resonance*, 19(3):222–241.

Chirapornchai, C., Bremner, P., and Daly, J. E. (2021). Helper's high with a robot pet. In *Companion of the 2021 ACM/IEEE International Conference on Human-Robot Interaction*, pages 229–233.

Chiu, C.-y., Dweck, C. S., Tong, J. Y.-y., and Fu, J. H.-y. (1997a). Implicit theories and conceptions of morality. *Journal of Personality and Social Psychology*, 73(5):923.

Chiu, C.-y., Hong, Y.-y., and Dweck, C. S. (1997b). Lay dispositionism and implicit theories of personality. *Journal of personality and social psychology*, 73(1):19.

Chmielewski, M. and Kucker, S. C. (2020). An mturk crisis? shifts in data quality and the impact on study results. *Social Psychological and Personality Science*, 11(4):464–473.

Cimpian, A., Arce, H.-M. C., Markman, E. M., and Dweck, C. S. (2007). Subtle linguistic cues affect children's motivation. *Psychological Science*, 18(4):314–316.

Clements, W. A. (2000). From an implicit to an explicit "theory of mind.". *Beyond dissociation: Interaction between dissociated implicit and explicit processing*, 22:273.

Clements, W. A. and Perner, J. (1994). Implicit understanding of belief. *Cognitive development*, 9(4):377–395.

Cobanoglu, C., Cavusoglu, M., and Turktarhan, G. (2021). A beginner's guide and best practices for using crowdsourcing platforms for survey research: The case of amazon mechanical turk (mturk). *Journal of Global Business Insights*, 6(1):92–97.

Cohen, J. (1992). Statistical power analysis. *Current Directions in Psychological Science*, 1(3):98–101.

Collins, E. C. (2019). Drawing parallels in human–other interactions: a trans-disciplinary approach to developing human–robot interaction methodologies. *Philosophical Transactions of the Royal Society B*, 374(1771):20180433.

Conger, S., Pratt, J. H., and Loch, K. D. (2013). Personal information privacy and emerging technologies. *Information Systems Journal*, 23(5):401–417.

Copleston, S. and Bugmann, G. (2008). Personal robot user expectations. *Advances in Communications, Computing, Networks and Security Volume*, 5:230.

Covington, M. V. and Omelich, C. L. (1985). Ability and effort valuation among failure-avoiding and failure-accepting students. *Journal of educational psychology*, 77(4):446.

Damholdt, M. F., Vestergaard, C., Nørskov, M., Hakli, R., Larsen, S., and Seibt, J. (2020). Towards a new scale for assessing attitudes towards social robots: The attitudes towards social robots scale (asor). *Interaction Studies*, 21(1):24–56.

Damiano, L. and Dumouchel, P. (2018). Anthropomorphism in human–robot co-evolution. *Frontiers in Psychology*, 9(468).

Dang, J. and Liu, L. (2021). Robots are friends as well as foes: Ambivalent attitudes toward mindful and mindless ai robots in the united states and china. *Computers in Human Behavior*, 115:106612.

Dautenhahn, K. (2007). Methodology & themes of human-robot interaction: A growing research field. *International Journal of Advanced Robotic Systems*, 4(1):15.

Dautenhahn, K., Woods, S., Kaouri, C., Walters, M. L., Koay, K. L., and Werry, I. (2005). What is a robot companion-friend, assistant or butler? In *2005 IEEE/RSJ international conference on intelligent robots and systems*, pages 1192–1197. IEEE.

Davison, D. P., Wijnen, F. M., Charisi, V., van der Meij, J., Reidsma, D., and Evers, V. (2021). Words of encouragement: how praise delivered by a social robot changes children's mindset for learning. *Journal on multimodal user interfaces*, 15(1):61–76.

De Castella, K. and Byrne, D. (2015). My intelligence may be more malleable than yours: the revised implicit theories of intelligence (self-theory) scale is a better predictor of achievement, motivation, and student disengagement. *European Journal of Psychology of Education*, 30(3):245–267.

De Castella, K., Goldin, P., Jazaieri, H., Ziv, M., Dweck, C. S., and Gross, J. J. (2013). Beliefs about emotion: Links to emotion regulation, well-being, and psychological distress. *Basic and Applied Social Psychology*, 35(6):497–505.

De Graaf, M. M. and Allouch, S. B. (2013). Exploring influencing variables for the acceptance of social robots. *Robotics and autonomous systems*, 61(12):1476–1486.

de Graaf, M. M. and Allouch, S. B. (2013). The relation between people's attitude and anxiety towards robots in human-robot interaction. In *2013 IEEE RO-MAN*, pages 632–637. IEEE.

De Graaf, M. M., Allouch, S. B., and Klamer, T. (2015). Sharing a life with harvey: Exploring the acceptance of and relationship-building with a social robot. *Computers in human behavior*, 43:1–14.

de Graaf, M. M., Allouch, S. B., and van Dijk, J. A. (2016). Long-term evaluation of a social robot in real homes. *Interaction studies*, 17(3):462–491.

de Graaf, M. M. A., Ben Allouch, S., and van Dijk, J. A. G. M. (2019). Why would i use this in my home? a model of domestic social robot acceptance. *Human–Computer Interaction*, 34(2):115–173.

de Kervenoael, R., Hasan, R., Schwob, A., and Goh, E. (2020). Leveraging human-robot interaction in hospitality services: Incorporating the role of perceived value, empathy, and information sharing into visitors' intentions to use social robots. *Tourism Management*, 78:104042.

de la Feria, R. and Grau Ruiz, M. A. (2022). Taxing robots. *Interactive Robotics: Legal, Ethical, Social and Economic Aspects (Springer Nature, 2022).*

Dennis, S. A., Goodson, B. M., and Pearson, C. A. (2020). Online worker fraud and evolving threats to the integrity of mturk data: A discussion of virtual private servers and the limitations of ip-based screening procedures. *Behavioral Research in Accounting*, 32(1):119–134.

Dinger, F. C. and Dickhäuser, O. (2013). Does implicit theory of intelligence cause achievement goals? evidence from an experimental study. *International Journal of Educational Research*, 61:38–47.

Doron, J., Stephan, Y., Boiché, J., and Scanff, C. L. (2009). Coping with examinations: Exploring relationships between students' coping strategies, implicit theories of ability, and perceived control. *British Journal of Educational Psychology*, 79(3):515–528.

Dudek, M., Baisch, S., Knopf, M., and Kolling, T. (2020). "this isn't me!": The role of age-related self- and user images for robot acceptance by elders. *International Journal of Social Robotics*.

Duffy, B. R. (2003). Anthropomorphism and the social robot. *Robotics and autonomous systems*, 42(3-4):177–190.

Dupeyrat, C. and Mariné, C. (2005). Implicit theories of intelligence, goal orientation, cognitive engagement, and achievement: A test of dweck's model with returning to school adults. *Contemporary educational psychology*, 30(1):43–59.

Dweck, C. (2017). *Mindset-updated edition: Changing the way you think to fulfil your potential*. Hachette UK.

Dweck, C. and Bempechat, J. (1983). Theories of intelligence and achievement motivation. *Learning and motivation in the classroom. Hillsdale, NJ: Erlbaum.*

Dweck, C. S. (1986). Motivational processes affecting learning. *American psychologist*, 41(10):1040.

Dweck, C. S. (2000). *Self-Theories: Their Role in Motivation, Personality, and Development. Essays in Social Psychology.* Psychology Press.

Dweck, C. S. (2008). Can personality be changed? the role of beliefs in personality and change. *Current directions in psychological science*, 17(6):391–394.

Dweck, C. S. (2009). Developing talent through a growth mindset. *Olympic Coach*, 21(1).

Dweck, C. S. (2013). *Self-theories: Their role in motivation, personality, and development*. Psychology press.

Dweck, C. S., Chiu, C.-y., and Hong, Y.-y. (1995a). Implicit theories and their role in judgments and reactions: A word from two perspectives. *Psychological inquiry*, 6(4):267–285.

Dweck, C. S., Chiu, C.-y., and Hong, Y.-y. (1995b). Implicit theories: Elaboration and extension of the model. *Psychological inquiry*, 6(4):322–333.

Dweck, C. S., Hong, Y. y., and Chiu, C. y. (1993). Implicit theories individual differences in the likelihood and meaning of dispositional inference. *Personality and Social Psychology Bulletin*, 19(5):644–656.

Dweck, C. S. and Leggett, E. L. (1988). A social-cognitive approach to motivation and personality. *Psychological review*, 95(2):256.

Dweck, C. S. and Sorich, L. (1999). Mastery-oriented thinking. *Coping*, 11:232–251.

Dweck, C. S. and Yeager, D. S. (2019). Mindsets: A view from two eras. *Perspectives on Psychological Science*, 14(3):481–496.

Dziergwa, M., Kaczmarek, M., Kaczmarek, P., Kedzierski, J., and Wadas-Szydlowska, K. (2018). Long-term cohabitation with a social robot: A case study of the influence of human attachment patterns. *International Journal of Social Robotics*, 10(1):163–176.

Einsiedel, E. F. (2009). *Emerging technologies: From hindsight to foresight*. UBC press.

Elliott, E. S. and Dweck, C. S. (1988). Goals: An approach to motivation and achievement. *Journal of personality and social psychology*, 54(1):5.

Enz, S., Diruf, M., Spielhagen, C., Zoll, C., and Vargas, P. A. (2011). The social role of robots in the future—explorative measurement of hopes and fears. *International Journal of Social Robotics*, 3(3):263.

Erdley, C. A. and Dweck, C. S. (1993). Children's implicit personality theories as predictors of their social judgments. *Child development*, 64(3):863–878.

Erdley, C. A., Loomis, C. C., Cain, K. M., and Dumas-Hines, F. (1997). Relations among children's social goals, implicit personality theories, and responses to social failure. *Developmental psychology*, 33(2):263.

Esterwood, C., Essenmacher, K., Yang, H., Zeng, F., and Robert, L. P. (2021). A meta-analysis of human personality and robot acceptance in human-robot interaction. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, pages 1–18. IEEE Press.

Esterwood, C. and Robert, L. P. (2022). Having the right attitude: How attitude impacts trust repair in human-robot interaction. In *Proceedings of the 2022 ACM/IEEE International Conference on Human-Robot Interaction*, page 332–341. IEEE Press.

Eurobarometer, S. (2012). Public attitudes towards robots. *Special Eurobarometer 382. European Commission.*

Eyssel, F. (2017). An experimental psychological perspective on social robotics. *Robotics and Autonomous Systems*, 87:363–371.

Eyssel, F., De Ruiter, L., Kuchenbrandt, D., Bobinger, S., and Hegel, F. (2012). 'if you sound like me, you must be more human': On the interplay of robot and user features on human-robot acceptance and anthropomorphism. In *2012 7th*

ACM/IEEE International Conference on Human-Robot Interaction (HRI), pages 125–126. IEEE.

Eyssel, F. and Hegel, F. (2012). (s) he's got the look: Gender stereotyping of robots 1. *Journal of Applied Social Psychology*, 42(9):2213–2230.

Ezer, N. (2008). *Is a robot an appliance, teammate, or friend? Age-related differences in expectations of and attitudes toward personal home-based robots.* Thesis, Georgia Institute of Technology.

Fasola, J. and Matarić, M. J. (2013). A socially assistive robot exercise coach for the elderly. *Journal of Human-Robot Interaction*, 2(2):3–32.

Faul, F., Erdfelder, E., Lang, A.-G., and Buchner, A. (2007). G*power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, 39(2):175–191.

Feil-Seifer, D., Haring, K. S., Rossi, S., Wagner, A. R., and Williams, T. (2020). Where to next? the impact of covid-19 on human-robot interaction research. *ACM Transactions on Human-Robot Interaction (THRI)*, 10(1):1–7.

Ferrari, F., Paladino, M. P., and Jetten, J. (2016). Blurring human–machine distinctions: anthropomorphic appearance in social robots as a threat to human distinctiveness. *International Journal of Social Robotics*, 8(2):287–302.

Floridi, L. (2017). Robots, jobs, taxes, and responsibilities. *Philosophy & Technology*, 30(1):1–4.

Fong, L. H. N., Chan, I. C. C., Law, R., and Ly, T. P. (2018). *The Mechanism that Links the Implicit Theories of Intelligence and Continuance of Information Technology: Evidence from the Use of Mobile Apps to Make Hotel Reservations*, pages 323–335. Springer.

Ford, J. B. (2017). Amazon's mechanical turk: a comment. *Journal of Advertising*, 46(1):156–158.

Fordsham, N., Moss, A. J., Krumholtz, S., Roggina, T., Robinson, J., and Litman, L. (2019). Variation among mechanical turk workers across time of day presents an opportunity and a challenge for research. *SAGE Open*.

Forgas-Coll, S., Huertas-Garcia, R., Andriella, A., and Alenyà, G. (2022). How do consumers' gender and rational thinking affect the acceptance of entertainment social robots? *International Journal of Social Robotics*, 14(4):973–994.

Freitas, A. L., Gollwitzer, P., and Trope, Y. (2004). The influence of abstract and concrete mindsets on anticipating and guiding others' self-regulatory efforts. *Journal of experimental social psychology*, 40(6):739–752.

Galin, R. and Mamchenko, M. (2021). Human-robot collaboration in the society of the future: A survey on the challenges and the barriers. In Singh, P. K., Veselov, G., Vyatkin, V., Pljonkin, A., Dodero, J. M., and Kumar, Y., editors, *Futuristic Trends in Network and Communication Technologies*, pages 111–122. Springer Singapore.

Gervey, B. M., Chiu, C.-y., Hong, Y.-y., and Dweck, C. S. (1999). Differential use of person information in decisions about guilt versus innocence: The role of implicit theories. *Personality and Social Psychology Bulletin*, 25(1):17–27.

Ghazali, A. S., Ham, J., Barakova, E., and Markopoulos, P. (2019). Assessing the effect of persuasive robots interactive social cues on users' psychological reactance, liking, trusting beliefs and compliance. *Advanced Robotics*, 33(7-8):325–337.

Ghazali, A. S., Ham, J., Barakova, E., and Markopoulos, P. (2020). Persuasive robots acceptance model (pram): roles of social responses within the acceptance model of persuasive robots. *International Journal of Social Robotics*, pages 1–18.

Ghiglino, D., De Tommaso, D., Willemse, C., Marchesi, S., and Wykowska, A. (2021). Can i get your (robot) attention? human sensitivity to subtle hints of human-likeness in a humanoid robot's behavior. *arXiv preprint*.

Glerum, J., Loyens, S. M. M., Wijnia, L., and Rikers, R. M. J. P. (2019). The effects of praise for effort versus praise for intelligence on vocational education students. *Educational Psychology*, pages 1–17.

Gnambs, T. and Appel, M. (2019). Are robots becoming unpopular? changes in attitudes towards autonomous robotic systems in europe. *Computers in Human Behavior*, 93:53–61.

Goetz, J., Kiesler, S., and Powers, A. (2003). Matching robot appearance and behavior to tasks to improve human-robot cooperation. In *The 12th IEEE International Workshop on Robot and Human Interactive Communication, 2003. Proceedings. ROMAN 2003.*, pages 55–60. Ieee.

Gollwitzer, P. M. (2012). *Mindset theory of action phases*, pages 526–545. Sage Publications Ltd, Thousand Oaks, CA.

Gouaillier, D., Hugel, V., Blazevic, P., Kilner, C., Monceaux, J., Lafourcade, P., Marnier, B., Serre, J., and Maisonnier, B. (2009). Mechatronic design of nao humanoid. In *2009 IEEE International Conference on Robotics and Automation*, pages 769–774. IEEE.

Grant, H. and Dweck, C. S. (2003). Clarifying achievement goals and their impact. *Journal of personality and social psychology*, 85(3):541.

Gray, K. and Wegner, D. M. (2012). Feeling robots and human zombies: Mind perception and the uncanny valley. *Cognition*, 125(1):125–130.

Greenwald, A. G., Banaji, M. R., Rudman, L. A., Farnham, S. D., Nosek, B. A., and Mellott, D. S. (2002). A unified theory of implicit attitudes, stereotypes, self-esteem, and self-concept. *Psychological review*, 109(1):3.

Guizzo, E. and Ackerman, E. (2012). How rethink robotics built its new baxter robot worker. *IEEE spectrum*, 7.

Gunderson, E. A., Gripshover, S. J., Romero, C., Dweck, C. S., Goldin-Meadow, S., and Levine, S. C. (2013). Parent praise to 1-to 3-year-olds predicts children's motivational frameworks 5 years later. *Child development*, 84(5):1526–1541.

Haegele, M. (2016). *World Robotics Service Robots, 2016*. IFR Statistical Department, VDMA.

Hafeez, A. (2019). *Promoting upskilling: How a situational growth mindset increases consumers' adoption of really new products*. Unpublished doctoral thesis, University of South-Eastern Norway.

Haimovitz, K. and Dweck, C. S. (2017). The origins of children's growth and fixed mindsets: New research and a new proposal. *Child Development*, 88(6):1849–1859.

Haimovitz, K. and Henderlong Corpus, J. (2011). Effects of person versus process praise on student motivation: Stability and change in emerging adulthood. *Educational Psychology*, 31(5):595–609.

Halperin, E., Russell, A. G., Trzesniewski, K. H., Gross, J. J., and Dweck, C. S. (2011). Promoting the middle east peace process by changing beliefs about group malleability. *Science*, 333(6050):1767–1769.

Ham, J. and Midden, C. J. (2014). A persuasive robot to stimulate energy conservation: the influence of positive and negative social feedback and task similarity on energy-consumption behavior. *International Journal of Social Robotics*, 6(2):163–171.

Han, B., Wang, L., and Li, X. (2019). To collaborate or serve? effects of anthropomorphized brand roles and implicit theories on consumer responses. *Cornell Hospitality Quarterly*, 61(1):53–67.

Haring, K. S., Silvera-Tawil, D., Matsumoto, Y., Velonaki, M., and Watanabe, K. (2014). Perception of an android robot in japan and australia: A cross-cultural comparison. In *International conference on social robotics*, pages 166–175. Springer.

Haring, K. S., Watanabe, K., Silvera-Tawil, D., Velonaki, M., and Takahashi, T. (2015). Changes in perception of a small humanoid robot. In *2015 6th International Conference on Automation, Robotics and Applications (ICARA)*, pages 83–89. IEEE.

Haring, K. S., Watanabe, K., Velonaki, M., Tossell, C. C., and Finomore, V. (2018). Ffab—the form function attribution bias in human–robot interaction. *IEEE Transactions on Cognitive and Developmental Systems*, 10(4):843–851.

Harrington, E. E., Bishop, A. J., Do, H. M., and Sheng, W. (2021). Perceptions of socially assistive robots: A pilot study exploring older adults' concerns. *Current Psychology*.

Hart, J., Hauser, E., Baker, S., Biswas, J., Jiao, J., and Sentis, L. (2022). Longitudinal social impacts of hri over long-term deployments. In *Proceedings of the 2022 ACM/IEEE International Conference on Human-Robot Interaction*, pages 1258–1260.

Haselhuhn, M. P., Schweitzer, M. E., and Kray, L. J. (2008). "beyond belief: How implicit beliefs influence trust". *Psychological Science*.

Haselhuhn, M. P., Schweitzer, M. E., Kray, L. J., and Kennedy, J. A. (2017). Perceptions of high integrity can persist after deception: How implicit beliefs moderate trust erosion. *Journal of Business Ethics*, 145(1):215–225.

Haselhuhn, M. P., Schweitzer, M. E., and Wood, A. M. (2010). How implicit beliefs influence trust recovery. *Psychological Science*, 21(5):645–648.

Haug, M. C. (2018). Fast, cheap, and unethical? the interplay of morality and methodology in crowdsourced survey research. *Review of Philosophy and Psychology*, 9(2):363–379.

Hauser, D., Paolacci, G., and Chandler, J. J. (2019). *Common concerns with MTurk as a participant pool: Evidence and solutions*, page 19. Routledge, New York.

He, Y., He, Q., and Liu, Q. (2022). Technology acceptance in socially assistive robots: Scoping review of models, measurement, and influencing factors. *Journal of Healthcare Engineering*, 2022:6334732.

Henrich, J., Heine, S. J., and Norenzayan, A. (2010). Most people are not weird. *Nature*, 466(7302):29–29.

Heslin, P. A., Latham, G. P., and VandeWalle, D. (2005). The effect of implicit person theory on performance appraisals. *Journal of Applied Psychology*, 90(5):842.

Hinks, T. (2020). Fear of robots and life satisfaction. *International Journal of Social Robotics*, pages 1–14.

Hoffman, G. and Zhao, X. (2020). A primer for conducting experiments in human–robot interaction. *J. Hum.-Robot Interact.*, 10(1):Article 6.

Hong, Y.-y., Chiu, C.-y., Dweck, C. S., Lin, D. M.-S., and Wan, W. (1999). Implicit theories, attributions, and coping: a meaning system approach. *Journal of Personality and Social psychology*, 77(3):588.

Hong, Y.-y., Chiu, C.-y., Dweck, C. S., and Sacks, R. (1997). Implicit theories and evaluative processes in person cognition. *Journal of Experimental Social Psychology*, 33(3):296–323.

Horowitz, M. C. (2016). Public opinion and the politics of the killer robots debate. *Research & Politics*, 3(1):2053168015627183.

Horstmann, A. C., Bock, N., Linhuber, E., Szczuka, J. M., Straßmann, C., and Krämer, N. C. (2018). Do a robot's social skills and its objection discourage interactants from switching the robot off? *PloS one*, 13(7):e0201581.

Hover, Q. R., Velner, E., Beelen, T., Boon, M., and Truong, K. P. (2021). Uncanny, sexy, and threatening robots: The online community's attitude to and perceptions of robots varying in humanlikeness and gender. In *Proceedings of the 2021 ACM/IEEE International Conference on Human-Robot Interaction*, pages 119–128.

Howell, A. J., Passmore, H.-A., and Holder, M. D. (2016). Implicit theories of well-being predict well-being and the endorsement of therapeutic lifestyle changes. *Journal of Happiness Studies*, 17(6):2347–2363.

Huang, H.-L., Cheng, L.-K., Sun, P.-C., and Chou, S.-J. (2021). The effects of perceived identity threat and realistic threat on the negative attitudes and usage intentions toward hotel service robots: The moderating effect of the robot's anthropomorphism. *International Journal of Social Robotics*.

Huang, N., Zuo, S., Wang, F., Cai, P., and Wang, F. (2017). The dark side of malleability: Incremental theory promotes immoral behaviors. *Frontiers in psychology*, 8:1341.

Huff, C. and Tingley, D. (2015). "who are these people?" evaluating the demographic characteristics and political preferences of mturk survey respondents. *Research & Politics*, 2(3):2053168015604648.

Hughes, J. S. (2015). Support for the domain specificity of implicit beliefs about persons, intelligence, and morality. *Personality and Individual Differences*, 86:195–203.

Irfan, B., Ramachandran, A., Spaulding, S., Parisi, G. I., and Gunes, H. (2022). Lifelong learning and personalization in long-term human-robot interaction (leap-hri). In *Proceedings of the 2022 ACM/IEEE International Conference on Human-Robot Interaction*, page 1261–1264.

Iwai, T. and de França Carvalho, J. V. (2020). Can a leopard change its spots? the effects of implicit theories of personality on forgiveness via attributions of behavioral stability. *Personality and Individual Differences*, 157:109800.

Jain, S. P., Mathur, P., and Maheswaran, D. (2009). The influence of consumers' lay theories on approach/avoidance motivation. *Journal of Marketing Research*, 46(1):56–65.

Jain, S. P. and Weiten, T. J. (2020). Consumer psychology of implicit theories: A review and agenda. *Consumer Psychology Review*, 3(1):60–75.

Jin, C.-H. (2013). The effects of individual innovativeness on users' adoption of internet content filtering software and attitudes toward children's internet use. *Computers in Human Behavior*, 29(5):1904–1916.

Johanson, G. A. and Brooks, G. P. (2010). Initial scale development: sample size for pilot studies. *Educational and psychological measurement*, 70(3):394–400.

Kahn, P. H., Reichert, A. L., Gary, H. E., Kanda, T., Ishiguro, H., Shen, S., Ruckert, J. H., and Gill, B. (2011). The new ontological category hypothesis in human-robot interaction. In *2011 6th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, pages 159–160. IEEE.

Kam, T.-K. (2011). Implicit theories and the trust repair process. In *22nd Annual IACM Conference Paper*.

Kamins, M. L. and Dweck, C. S. (1999). Person versus process praise and criticism: Implications for contingent self-worth and coping. *Developmental psychology*, 35(3):835.

Kammrath, L. K. and Peetz, J. (2012). You promised you'd change: How incremental and entity theorists react to a romantic partner's promised change attempts. *Journal of Experimental Social Psychology*, 48(2):570–574.

Kanero, J., Oranç, C., Koşkulu, S., Kumkale, G. T., Göksun, T., and Küntay, A. C. (2021). Are tutor robots for everyone? the influence of attitudes, anxiety, and personality on robot-led language learning. *International Journal of Social Robotics*, pages 1–16.

Kees, J., Berry, C., Burton, S., and Sheehan, K. (2017). An analysis of data quality: Professional panels, student subject pools, and amazon's mechanical turk. *Journal of Advertising*, 46(1):141–155.

Keizer, R. A. O., Van Velsen, L., Moncharmont, M., Riche, B., Ammour, N., Del Signore, S., Zia, G., Hermens, H., and N'Dja, A. (2019). Using socially assistive robots for monitoring and preventing frailty among older adults: A study on usability and user experience challenges. *Health and Technology*, 9(4):595–605.

Kelley, K. and Maxwell, S. E. (2003). Sample size for multiple regression: Obtaining regression coefficients that are accurate, not simply significant. *Psychological Methods*, 8:305–321.

Kieslich, K., Lünich, M., and Marcinkowski, F. (2021). The threats of artificial intelligence scale (tai). *International Journal of Social Robotics*.

Kim, H. C. and Kramer, T. (2015). Do materialists prefer the "brand-as-servant"? the interactive effect of anthropomorphized brand roles and materialism on consumer responses. *Journal of Consumer Research*, 42(2):284–299.

Kim, S., Zhang, K., and Park, D. (2018). Don't want to look dumb? the role of theories of intelligence and humanlike features in online help seeking. *Psychological science*, 29(2):171–180.

King, R. B. (2016). A fixed mindset leads to negative affect. *Zeitschrift für Psychologie*, 225(2):137–145.

King, R. B. and dela Rosa, E. D. (2019). Are your emotions under your control or not? implicit theories of emotion predict well-being via cognitive reappraisal. *Personality and Individual Differences*, 138:177–182.

King, R. B., McInerney, D. M., and Watkins, D. A. (2012). How you think about your intelligence determines how you feel in school: The role of theories of intelligence on academic emotions. *Learning and Individual Differences*, 22(6):814–819.

Knee, C. R. (1998). Implicit theories of relationships: Assessment and prediction of romantic relationship initiation, coping, and longevity. *Journal of Personality and Social Psychology*, 74(2):360.

Knee, C. R., Patrick, H., and Lonsbary, C. (2003). Implicit theories of relationships: Orientations toward evaluation and cultivation. *Personality and Social Psychology Review*, 7(1):41–55.

Knee, C. R., Patrick, H., Vietor, N. A., and Neighbors, C. (2004). Implicit theories of relationships: Moderators of the link between conflict and commitment. *Personality and Social Psychology Bulletin*, 30(5):617–628.

Kok, B. C. and Soh, H. (2020). Trust in robots: Challenges and opportunities. *Current Robotics Reports*, 1(4):297–309.

Krägeloh, C. U., Bharatharaj, J., Sasthan Kutty, S. K., Nirmala, P. R., and Huang, L. (2019). Questionnaires to measure acceptability of social robots: A critical review. *Robotics*, 8(4):88.

Kuo, I. H., Rabindran, J. M., Broadbent, E., Lee, Y. I., Kerse, N., Stafford, R., and MacDonald, B. A. (2009). Age and gender factors in user acceptance of healthcare robots. In *RO-MAN 2009-The 18th IEEE International Symposium on Robot and Human Interactive Communication*, pages 214–219. IEEE.

Kühnlenz, B., Kühnlenz, K., Busse, F., Förtsch, P., and Wolf, M. (2018). Effect of explicit emotional adaptation on prosocial behavior of humans towards robots depends on prior robot experience. In *2018 27th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*, pages 275–281. IEEE.

Lam, S.-f., Yim, P.-s., and Ng, Y.-l. (2008). Is effort praise motivational? the role of beliefs in the effort–ability relationship. *Contemporary Educational Psychology*, 33(4):694–710.

Latikka, R., Savela, N., Koivula, A., and Oksanen, A. (2021). Attitudes toward robots as equipment and coworkers and the impact of robot autonomy level. *International Journal of Social Robotics*.

Latikka, R., Turja, T., and Oksanen, A. (2019). Self-efficacy and acceptance of robots. *Computers in Human Behavior*, 93:157–163.

Lee, C.-S., Wang, M.-H., Yen, S.-J., Wei, T.-H., Wu, I.-C., Chou, P.-C., Chou, C.-H., Wang, M.-W., and Yan, T.-H. (2016). Human vs. computer go: Review and prospect [discussion forum]. *IEEE Computational Intelligence Magazine*, 11(3):67–72.

Lee, H. R. and Šabanović, S. (2014). Culturally variable preferences for robot design and use in south korea, turkey, and the united states. In *2014 9th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, pages 17–24. IEEE.

Lee, K. M., Peng, W., Jin, S.-A., and Yan, C. (2006). Can robots manifest personality?: An empirical test of personality recognition, social responses, and social presence in human–robot interaction. *Journal of communication*, 56(4):754–772.

Leite, I., Castellano, G., Pereira, A., Martinho, C., and Paiva, A. (2012). Long-term interactions with empathic robots: Evaluating perceived support in children. In *International Conference on Social Robotics*, pages 298–307. Springer.

Leoste, J., Viik, T., López, J. S. M., Kangur, M., Vunder, V., Mollard, Y., Õun, T., Tammo, H., and Paekivi, K. (2021). Robots as my future colleagues: Changing attitudes toward collaborative robots by means of experience-based workshops. In Mealha, , Dascalu, M., and Di Mascio, T., editors, *Ludic, Co-design and Tools Supporting Smart Learning Ecosystems and Smart Education*, pages 163–175. Springer Singapore.

Levy, S. R. and Dweck, C. S. (1998). Trait-versus process-focused social judgment. *Social Cognition*, 16(1):151–172.

Levy, S. R., Stroessner, S. J., and Dweck, C. S. (1998). Stereotype formation and endorsement: The role of implicit theories. *Journal of Personality and Social Psychology*, 74(6):1421.

Lewis, M., Sycara, K., and Walker, P. (2018). *The role of trust in human-robot interaction*, pages 135–159. Springer, Cham.

Leyzberg, D., Spaulding, S., and Scassellati, B. (2014). Personalizing robot tutors to individuals' learning differences. In *2014 9th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, pages 423–430. IEEE.

Li, F., Harmer, P., Duncan, T. E., Duncan, S. C., Acock, A., and Boles, S. (1998). Approaches to testing interaction effects using structural equation modeling methodology. *Multivariate Behavioral Research*, 33(1):1–39.

Li, Y. and Bates, T. C. (2019). You can't change your basic ability, but you work at things, and that's how we get hard things done: Testing the role of growth mindset on response to setbacks, educational attainment, and cognitive ability. *Journal of Experimental Psychology: General*, 148(9):1640.

Liang, Y. and Lee, S. A. (2017). Fear of autonomous robots and artificial intelligence: Evidence from national representative data with probability sampling. *International Journal of Social Robotics*, 9(3):379–384.

Lin, C., Šabanović, S., Dombrowski, L., Miller, A. D., Brady, E., and MacDorman, K. F. (2021). Parental acceptance of children's storytelling robots: A projection of the uncanny valley of ai. *Frontiers in Robotics and AI*, 8.

Linxen, S., Sturm, C., Brühlmann, F., Cassau, V., Opwis, K., and Reinecke, K. (2021). How weird is chi? In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, pages 1–14.

Lou, N. M. and Noels, K. A. (2020). "does my teacher believe i can improve?": The role of meta-lay theories in esl learners' mindsets and need satisfaction. *Frontiers in psychology*, 11:1417.

Lu, L., Neale, N., Line, N. D., and Bonn, M. (2021). Improving data quality using amazon mechanical turk through platform setup. *Cornell Hospitality Quarterly*, page 19389655211025475.

Mangels, J. A., Butterfield, B., Lamb, J., Good, C., and Dweck, C. S. (2006). Why do beliefs about intelligence influence learning success? a social cognitive neuroscience model. *Social cognitive and affective neuroscience*, 1(2):75–86.

Manzi, F., Massaro, D., Di Lernia, D., Maggioni, M. A., Riva, G., and Marchetti, A. (2021). Robots are not all the same: Young adults' expectations, attitudes, and mental attribution to two humanoid social robots. *Cyberpsychology, Behavior, and Social Networking*, 24(5):307–314.

Mathur, P., Block, L., and Yucel-Aybat, O. (2014). The effects of goal progress cues: An implicit theory perspective. *Journal of Consumer Psychology*, 24(4):484–496.

Mathur, P., Chun, H. H., and Maheswaran, D. (2016). Consumer mindsets and self-enhancement: Signaling versus learning. *Journal of Consumer Psychology*, 26(1):142–152.

Mathur, P., Jain, S. P., Hsieh, M.-H., Lindsey, C. D., and Maheswaran, D. (2013). The influence of implicit theories and message frame on the persuasiveness of disease prevention and detection advocacies. *Organizational Behavior and Human Decision Processes*, 122(2):141–151.

Mathur, P., Jain, S. P., and Maheswaran, D. (2012). Consumers' implicit theories about personality influence their brand personality judgments. *Journal of Consumer Psychology*, 22(4):545–557.

Matthews, G., Hancock, P. A., Lin, J., Panganiban, A. R., Reinerman-Jones, L. E., Szalma, J. L., and Wohleber, R. W. (2020). Evolution and revolution: Personality research for the coming world of robots, artificial intelligence, and autonomous systems. *Personality and Individual Differences*, page 109969.

Maxwell, S. E., Kelley, K., and Rausch, J. R. (2008). Sample size planning for statistical power and accuracy in parameter estimation. *Annual Review of Psychology*, 59(1):537–563.

McConnell, A. R. (2001). Implicit theories: Consequences for social judgments of individuals. *Journal of Experimental Social Psychology*, 37(3):215–227.

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

Meyers, E. A., Walker, A. C., Fugelsang, J. A., and Koehler, D. J. (2020). Reducing the number of non-naïve participants in mechanical turk samples. *Methods in Psychology*, 3:100032.

Midden, C. and Ham, J. (2009). Using negative and positive social feedback from a robotic agent to save energy. In *Proceedings of the 4th international conference on persuasive technology*, page 12. ACM.

Miller, C. H., Burgoon, J. K., and Hall, J. R. (2007). The effects of implicit theories of moral character on affective reactions to moral transgressions. *Social Cognition*, 25(6):819–832.

Mirsky, R., Zimmerman, M., Ahmad, M., Bagchi, S., Gervits, F., Han, Z., Hart, J., García, D. H., Leonetti, M., Mead, R., Senft, E., Sinapov, J., and Wilson, J. (2021). Ai-hri 2021 proceedings.

Miu, A. S. and Yeager, D. S. (2015). Preventing symptoms of depression by teaching adolescents that people can change: Effects of a brief incremental theory of personality intervention at 9-month follow-up. *Clinical Psychological Science*, 3(5):726–743.

Molden, D. C. and Dweck, C. S. (2006). Finding" meaning" in psychology: a lay theories approach to self-regulation, social perception, and social development. *American Psychologist*, 61(3):192.

Montford, W. J., Leary, R. B., and Nagel, D. M. (2019). The impact of implicit self-theories and loss salience on financial risk. *Journal of Business Research*, 99:1–11.

Morsunbul, U. (2019). Human-robot interaction: How do personality traits affect attitudes towards robot? *Journal of Human Sciences*, 16(2):499–504.

Moser, J. S., Schroder, H. S., Heeter, C., Moran, T. P., and Lee, Y.-H. (2011). Mind your errors: Evidence for a neural mechanism linking growth mind-set to adaptive posterror adjustments. *Psychological Science*, 22(12):1484–1489.

Mueller, C. M. and Dweck, C. S. (1998). Praise for intelligence can undermine children's motivation and performance. *Journal of personality and social psychology*, 75(1):33.

Munger, K., Luca, M., Nagler, J., and Tucker, J. (2018). Everyone on mechanical turk is above a threshold of digital literacy: Sampling strategies for studying digital media effects. Report, Working Paper. https://csdp. princeton. edu/sites/csdp/files/media/munger ....

Murphy, M. C. and Dweck, C. S. (2016a). Mindsets and consumer psychology: A response. *Journal of Consumer Psychology*, 26(1):165–166.

Murphy, M. C. and Dweck, C. S. (2016b). Mindsets shape consumer behavior. *Journal of Consumer Psychology*, 26(1):127–136.

Mutlu, B., Osman, S., Forlizzi, J., Hodgins, J., and Kiesler, S. (2006). Task structure and user attributes as elements of human-robot interaction design. In *ROMAN 2006-The 15th IEEE International Symposium on Robot and Human Interactive Communication*, pages 74–79. IEEE.

Naneva, S., Sarda Gou, M., Webb, T. L., and Prescott, T. J. (2020). A systematic review of attitudes, anxiety, acceptance, and trust towards social robots. *International Journal of Social Robotics*, 12(6):1179–1201.

Nash, K., Lea, J. M., Davies, T., and Yogeeswaran, K. (2018). The bionic blues: Robot rejection lowers self-esteem. *Computers in human behavior*, 78:59–63bionic b.

Neyer, F. J., Felber, J., and Gebhardt, C. (2012). Entwicklung und validierung einer kurzskala zur erfassung von technikbereitschaft. *Diagnostica*.

Ng, A. S. and Tong, E. M. (2013). The relation between implicit theories of personality and forgiveness. *Personal Relationships*, 20(3):479–494.

Ninomiya, T., Fujita, A., Suzuki, D., and Umemuro, H. (2015). Development of the multi-dimensional robot attitude scale: Constructs of people's attitudes towards domestic robots. In *International Conference on Social Robotics*, pages 482–491. Springer.

Nomura, T., Kanda, T., Kidokoro, H., Suehiro, Y., and Yamada, S. (2016). Why do children abuse robots? *Interaction Studies*, 17(3):347–369.

Nomura, T., Kanda, T., Suzuki, T., and Kato, K. (2004). Psychology in human-robot communication: An attempt through investigation of negative attitudes and anxiety toward robots. In *RO-MAN 2004. 13th IEEE International Workshop on Robot and Human Interactive Communication (IEEE Catalog No. 04TH8759)*, pages 35–40. IEEE.

Nomura, T., Kanda, T., Suzuki, T., and Kato, K. (2009). Age differences and images of robots: Social survey in japan. *Interaction Studies*, 10(3):374–391.

Nomura, T. and Tanaka, M. (2022). Experiences, knowledge of functions, and social acceptance of robots: an exploratory case study focusing on japan. *AI SOCIETY*, 37(1):367–374.

Nussbaum, A. D. and Dweck, C. S. (2008). Defensiveness versus remediation: Self-theories and modes of self-esteem maintenance. *Personality and Social Psychology Bulletin*, 34(5):599–612.

Oh, C., Lee, T., Kim, Y., Park, S., Kwon, S., and Suh, B. (2017). Us vs. them: Understanding artificial intelligence technophobia over the google deepmind challenge match. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, CHI '17, pages 2523–2534, New York, NY, USA. ACM.

Ommundsen, Y., Haugen, R., and Lund, T. (2005). Academic self-concept, implicit theories of ability, and self-regulation strategies. *Scandinavian Journal of Educational Research*, 49(5):461–474.

Ostrowski, A. K., Breazeal, C., and Park, H. W. (2022). Mixed-method long-term robot usage: Older adults' lived experience of social robots. In *Proceedings of the 2022 ACM/IEEE International Conference on Human-Robot Interaction*, page 33–42.

Park, H. W., Rosenberg-Kima, R., Rosenberg, M., Gordon, G., and Breazeal, C. (2017). Growing growth mindset with a social robot peer. In *2017 12th ACM/IEEE International Conference on Human-Robot Interaction (HRI*, pages 137–145. IEEE.

Park, J. K. and John, D. R. (2010). Got to get you into my life: Do brand personalities rub off on consumers? *Journal of consumer research*, 37(4):655–669.

Park, J. K. and John, D. R. (2012). Capitalizing on brand personalities in advertising: The influence of implicit self-theories on ad appeal effectiveness. *Journal of Consumer Psychology*, 22(3):424–432.

Park, J. K. and John, D. R. (2018). Developing brand relationships after a brand transgression: The role of implicit theories of relationships. *Journal of the Association for Consumer Research*, 3(2):175–187.

Paunesku, D. (2013). *Scaled-up social psychology: Intervening wisely and broadly in education*. Unpublished doctoral thesis, Stanford University Stanford, CA.

Paunesku, D., Walton, G. M., Romero, C., Smith, E. N., Yeager, D. S., and Dweck, C. S. (2015). Mind-set interventions are a scalable treatment for academic underachievement. *Psychological science*, 26(6):784–793.

Peer, E., Rothschild, D. M., Evernden, Z., Gordon, A., and Damer, E. (2021). Mturk, prolific or panels? choosing the right audience for online research. *Choosing the right audience for online research (January 10, 2021)*.

Petrocelli, J. V., Clarkson, J. J., Tormala, Z. L., and Hendrix, K. S. (2010). Perceiving stability as a means to attitude certainty: The role of implicit theories of attitudes. *Journal of Experimental Social Psychology*, 46(6):874–883.

Plaks, J. E., Grant, H., and Dweck, C. S. (2005). Violations of implicit theories and the sense of prediction and control: implications for motivated person perception. *Journal of Personality and Social Psychology*, 88(2):245.

Plaks, J. E. and Stecher, K. (2007). Unexpected improvement, decline, and stasis: A prediction confidence perspective on achievement success and failure. *Journal of personality and social psychology*, 93(4):667.

Plaks, J. E., Stroessner, S. J., Dweck, C. S., and Sherman, J. W. (2001). Person theories and attention allocation: Preferences for stereotypic versus counterstereotypic information. *Journal of personality and social psychology*, 80(6):876.

Podsakoff, P. M., MacKenzie, S. B., Lee, J.-Y., and Podsakoff, N. P. (2003). Common method biases in behavioral research: a critical review of the literature and recommended remedies. *Journal of applied psychology*, 88(5):879.

Pomerantz, E. M. and Kempner, S. G. (2013). Mothers' daily person and process praise: Implications for children's theory of intelligence and motivation. *Developmental Psychology*, 49(11):2040.

Poon, C. S. and Koehler, D. J. (2006). Lay personality knowledge and dispositionist thinking: A knowledge-activation framework. *Journal of Experimental Social Psychology*, 42(2):177–191.

Price, L. L., Coulter, R. A., Strizhakova, Y., and Schultz, A. E. (2017). The fresh start mindset: transforming consumers' lives. *Journal of Consumer Research*, 45(1):21–48.

Priester, J. R. and Petty, R. E. (2016). A research dialogue on mindsets. *Journal of Consumer Psychology*, 26(1):125–126.

Pyo, J. and Maxfield, M. G. (2021). Cognitive effects of inattentive responding in an mturk sample. *Social Science Quarterly*.

Quintanilla, V. D. (2011). Judicial mindsets: The social psychology of implicit theories and the law. *Neb. L. Rev.*, 90:611.

Rai, D. and Lin, C.-W. W. (2019). The influence of implicit self-theories on consumer financial decision making. *Journal of Business Research*, 95:316–325.

Rattan, A., Good, C., and Dweck, C. S. (2012). "it's ok—not everyone can be good at math": Instructors with an entity theory comfort (and demotivate) students. *Journal of Experimental Social Psychology*, 48(3):731–737.

Rauthmann, J. F. (2020). Personality science: United in diversity. *Personality Science*, 1(1):1–25.

Ray, C., Mondada, F., and Siegwart, R. (2008). What do people expect from robots? In *2008 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 3816–3821. IEEE.

Rea, D. J., Schneider, S., and Kanda, T. (2021). " is this all you can do? harder!" the effects of (im) polite robot encouragement on exercise effort. In *Proceedings of the 2021 ACM/IEEE International Conference on Human-Robot Interaction*, pages 225–233.

Reich-Stiebert, N. and Eyssel, F. (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):875–888.

Reich-Stiebert, N., Eyssel, F., and Hohnemann, C. (2019). Involve the user! changing attitudes toward robots by user participation in a robot prototyping process. *Computers in Human Behavior*, 91:290–296.

Rhodewalt, F. (1994). Conceptions of ability, achievement goals, and individual differences in self-handicapping behavior: On the application of implicit theories. *Journal of personality*, 62(1):67–85.

Ricci-Cabello, I., Bobrow, K., Islam, S. M. S., Chow, C. K., Maddison, R., Whittaker, R., and Farmer, A. J. (2019). Examining development processes for text messaging interventions to prevent cardiovascular disease: systematic literature review. *JMIR mHealth and uHealth*, 7(3):e12191.

Robert, L., Alahmad, R., Esterwood, C., Kim, S., You, S., and Zhang, Q. (2020). A review of personality in human–robot interactions. *Available at SSRN 3528496.*

Robinette, P., Howard, A. M., and Wagner, A. R. (2017). Effect of robot performance on human–robot trust in time-critical situations. *IEEE Transactions on Human-Machine Systems*, 47(4):425–436.

Robins, R. W. and Pals, J. L. (2002). Implicit self-theories in the academic domain: Implications for goal orientation, attributions, affect, and self-esteem change. *Self and identity*, 1(4):313–336.

Robinson, J., Rosenzweig, C., Moss, A. J., and Litman, L. (2019). Tapped out or barely tapped? recommendations for how to harness the vast and largely unused potential of the mechanical turk participant pool. *PloS one*, 14(12).

Rossi, A., Dautenhahn, K., Koay, K. L., and Saunders, J. (2017). Investigating human perceptions of trust in robots for safe hri in home environments. In *Proceedings of the Companion of the 2017 ACM/IEEE International Conference on Human-Robot Interaction*, pages 375–376.

Rossi, A., Dautenhahn, K., Koay, K. L., and Walters, M. L. (2018). The impact of peoples' personal dispositions and personalities on their trust of robots in an emergency scenario. *Paladyn, Journal of Behavioral Robotics*, 9(1):137–154.

Rossi, A., Dautenhahn, K., Koay, K. L., Walters, M. L., and Holthaus, P. (2020). Evaluating people's perceptions of trust in a robot in a repeated interactions study. In *International Conference on Social Robotics*, pages 453–465. Springer.

Rucker, D. D. and Galinsky, A. D. (2016). Growing beyond growth: Why multiple mindsets matter for consumer behavior. *Journal of Consumer Psychology*, 26(1):161–164.

Rucker, D. D. and He, S. (2016). Psychological mindsets affect consumption: How different mindsets help (hurt) portion control. *Appetite*, 103:425–431.

Rueben, M., Nikolaidis, S., de Graaf, M., Phillips, E., Robert, L., Sirkin, D., Kwon, M., and Thellman, S. (2020). Half day workshop on mental models of robots. In *Companion of the 2020 ACM/IEEE International Conference on Human-Robot Interaction*, pages 658–659.

Rydell, R. J., Hugenberg, K., Ray, D., and Mackie, D. M. (2007). Implicit theories about groups and stereotyping: The role of group entitativity. *Personality and Social Psychology Bulletin*, 33(4):549–558.

Saari, U. A., Tossavainen, A., Kaipainen, K., and Mäkinen, S. J. (2022). Exploring factors influencing the acceptance of social robots among early adopters and mass market representatives. *Robotics and Autonomous Systems*, 151:104033.

Salvini, P., Ciaravella, G., Yu, W., Ferri, G., Manzi, A., Mazzolai, B., Laschi, C., Oh, S.-R., and Dario, P. (2010). How safe are service robots in urban environments? bullying a robot. In *19th international symposium in robot and human interactive communication*, pages 1–7. IEEE.

Sanders, T. L., MacArthur, K., Volante, W., Hancock, G., MacGillivray, T., Shugars, W., and Hancock, P. (2017). Trust and prior experience in human-robot interaction. In *Proceedings of the human factors and ergonomics society annual meeting*, volume 61, pages 1809–1813. SAGE Publications Sage CA: Los Angeles, CA.

Sandoval, E. B., Mubin, O., and Obaid, M. (2014). Human robot interaction and fiction: A contradiction. In *International Conference on Social Robotics*, pages 54–63. Springer.

Santamaria, T. and Nathan-Roberts, D. (2017). Personality measurement and design in human-robot interaction: A systematic and critical review. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, volume 61, pages 853–857. SAGE Publications Sage CA: Los Angeles, CA.

Savela, N., Turja, T., and Oksanen, A. (2018). Social acceptance of robots in different occupational fields: A systematic literature review. *International Journal of Social Robotics*, 10(4):493–502.

Schaefer, K. E. (2016). *Measuring trust in human robot interactions: Development of the "trust perception scale-HRI"*, pages 191–218. Springer.

Schepers, J. and Streukens, S. (2022). To serve and protect: a typology of service robots and their role in physically safe services. *Journal of Service Management*, 33(2):197–209.

Schermerhorn, P., Scheutz, M., and Crowell, C. R. (2008). Robot social presence and gender: Do females view robots differently than males? In *Proceedings of the 3rd ACM/IEEE international conference on Human robot interaction*, pages 263–270. ACM.

Schleider, J. and Schroder, H. (2018). Implicit theories of personality across development: Impacts on coping, resilience, and mental health. *The SAGE Handbook of Personality and Individual Differences. Sage Publications*.

Schleider, J. L. and Weisz, J. R. (2016). Reducing risk for anxiety and depression in adolescents: Effects of a single-session intervention teaching that personality can change. *Behaviour research and therapy*, 87:170–181.

Schneider, S. and Kummert, F. (2016). Motivational effects of acknowledging feedback from a socially assistive robot. In *International Conference on Social Robotics*, pages 870–879. Springer.

Schroder, H. S., Dawood, S., Yalch, M. M., Donnellan, M. B., and Moser, J. S. (2016). Evaluating the domain specificity of mental health–related mind-sets. *Social Psychological and Personality Science*, 7(6):508–520.

Schroder, H. S., Moran, T. P., Donnellan, M. B., and Moser, J. S. (2014). Mindset induction effects on cognitive control: A neurobehavioral investigation. *Biological psychology*, 103:27–37.

Schüle, M., Kraus, J. M., Babel, F., and Reißner, N. (2022). Patients' trust in hospital transport robots: Evaluation of the role of user dispositions, anxiety, and robot characteristics. In *Proceedings of the 2022 ACM/IEEE International Conference on Human-Robot Interaction*, pages 246–255.

Schumann, K. and Dweck, C. S. (2014). Who accepts responsibility for their transgressions? *Personality and social psychology bulletin*, 40(12):1598–1610.

Schönbrodt, F. D. and Perugini, M. (2013). At what sample size do correlations stabilize? *Journal of Research in Personality*, 47(5):609–612.

Seo, Y., Ko, D., and Kim, J. (2021). It is all in the mind (set)! matching mindsets and luxury tourism. *Journal of Travel Research*, 60(1):184–196.

Septianto, F., Seo, Y., and Paramita, W. (2021). Express: The role of implicit theories in motivating donations in response to threat-based awe. *Journal of Public Policy & Marketing*, page 07439156211042281.

Serholt, S. and Barendregt, W. (2016). Robots tutoring children: Longitudinal evaluation of social engagement in child-robot interaction. In *Proceedings of the 9th nordic conference on human-computer interaction*, pages 1–10.

Severson, R. L. and Carlson, S. M. (2010). Behaving as or behaving as if? children's conceptions of personified robots and the emergence of a new ontological category. *Neural Networks*, 23(8-9):1099–1103.

Sevincer, A. T., Kluge, L., and Oettingen, G. (2014). Implicit theories and motivational focus: Desired future versus present reality. *Motivation and Emotion*, 38(1):36–46.

Sharifi, S. S. and Palmeira, M. (2017). Customers' reactions to technological products: The impact of implicit theories of intelligence. *Computers in Human Behavior*, 77:309–316.

Skipper, Y. and Douglas, K. (2012). Is no praise good praise? effects of positive feedback on children's and university students' responses to subsequent failures. *British Journal of Educational Psychology*, 82(2):327–339.

Smith, E. R., Šabanović, S., and Fraune, M. R. (2021). Human-robot interaction through the lens of social psychological theories of intergroup behavior. *Technology, Mind, and Behavior*, 1.

Solberg, E., Traavik, L. E., and Wong, S. I. (2020). Digital mindsets: Recognizing and leveraging individual beliefs for digital transformation. *California Management Review*, page 0008125620931839.

Song, Y.-A., Lee, S. Y., and Kim, Y. (2019). Does mindset matter for using social networking sites?: understanding motivations for and uses of instagram with growth versus fixed mindset. *International Journal of Advertising*, pages 1–19.

Spatola, N., Kühnlenz, B., and Cheng, G. (2021). Perception and evaluation in human–robot interaction: The human–robot interaction evaluation scale (hries)—a multicomponent approach of anthropomorphism. *International Journal of Social Robotics*, 13(7):1517–1539.

Spatola, N. and Wudarczyk, O. A. (2021). Implicit attitudes towards robots predict explicit attitudes, semantic distance between robots and humans, anthropomorphism, and prosocial behavior: From attitudes to human–robot interaction. *International Journal of Social Robotics*, 13(5):1149–1159.

Spinath, B., Spinath, F. M., Riemann, R., and Angleitner, A. (2003). Implicit theories about personality and intelligence and their relationship to actual personality and intelligence. *Personality and Individual Differences*, 35(4):939–951.

Stapels, J. G. and Eyssel, F. (2021). Robocalypse? yes, please! the role of robot autonomy in the development of ambivalent attitudes towards robots. *International Journal of Social Robotics*, pages 1–15.

Stock-Homburg, R., Kirchhoff, J., Heinisch, J. S., Ebert, A., Busch, P., Rawal, N., David, K., Wendt, J., Spiecker gen Döhmann, I., and von Stryk, O. (2022). Responsible human-robot interaction with anthropomorphic service robots: State of the art of an interdisciplinary research challenge. *Publications of Darmstadt Technical University, Institute for Business Studies (BWL)*.

Strait, M., Urry, H. L., and Muentener, P. (2019). Children's responding to humanlike agents reflects an uncanny valley. In *2019 14th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, pages 506–515. IEEE.

Strait, M. K., Aguillon, C., Contreras, V., and Garcia, N. (2017). The public's perception of humanlike robots: Online social commentary reflects an appearance-based uncanny valley, a general fear of a "technology takeover", and the unabashed sexualization of female-gendered robots. In *2017 26th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*, pages 1418–1423. IEEE.

Swift-Spong, K., Short, E., Wade, E., and Matarić, M. J. (2015). Effects of comparative feedback from a socially assistive robot on self-efficacy in post-stroke rehabilitation. In *2015 IEEE International Conference on Rehabilitation Robotics (ICORR)*, pages 764–769. IEEE.

Syrdal, D. S., Dautenhahn, K., Koay, K. L., and Walters, M. L. (2009). The negative attitudes towards robots scale and reactions to robot behaviour in a live human-robot interaction study. *Adaptive and emergent behaviour and complex systems*.

Takayama, L., Ju, W., and Nass, C. (2008). Beyond dirty, dangerous and dull: what everyday people think robots should do. In *2008 3rd ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, pages 25–32. IEEE.

Tam, K., Shu, T., Ng, H. K., and Tong, Y. (2013). Belief about immutability of moral character and punitiveness toward criminal offenders. *Journal of Applied Social Psychology*, 43(3):603–611.

Tamir, M., John, O. P., Srivastava, S., and Gross, J. J. (2007). Implicit theories of emotion: Affective and social outcomes across a major life transition. *Journal of personality and social psychology*, 92(4):731.

Tan, X. Z., Vázquez, M., Carter, E. J., Morales, C. G., and Steinfeld, A. (2018). Inducing bystander interventions during robot abuse with social mechanisms. In *Proceedings of the 2018 ACM/IEEE international conference on human-robot interaction*, pages 169–177.

Tapus, A., Ţăpuş, C., and Matarić, M. J. (2008). User—robot personality matching and assistive robot behavior adaptation for post-stroke rehabilitation therapy. *Intelligent Service Robotics*, 1(2):169–183.

Tavakol, M. and Dennick, R. (2011). Making sense of cronbach's alpha. *Int J Med Educ*, 2:53–55.

Thompson, C., Mohamed, S., Louie, W.-Y. G., He, J. C., Li, J., and Nejat, G. (2017). The robot tangy facilitating trivia games: A team-based user-study with long-term care residents. In *2017 IEEE international symposium on robotics and intelligent sensors (IRIS)*, pages 173–178. IEEE.

Troncone, A., Amorese, T., Cuciniello, M., Saturno, R., Pugliese, L., Cordasco, G., Vogel, C., and Esposito, A. (2020). Advanced assistive technologies for elderly people: a psychological perspective on seniors' needs and preferences (part a). *Acta Polytechnica Hungarica*, 17(2):163–189.

Turja, T. and Oksanen, A. (2019). Robot acceptance at work: A multilevel analysis based on 27 eu countries. *International Journal of Social Robotics*, 11(4):679–689.

Tzeng, J.-Y. and Chen, C.-T. (2012). Computer praise, attributional orientations, and games: A reexamination of the casa theory relative to children. *Computers in Human Behavior*, 28(6):2420–2430.

van Wynsberghe, A., Ley, M., and Roeser, S. (2022). *Ethical Aspects of Human–Robot Collaboration in Industrial Work Settings*, pages 255–266. Springer International Publishing, Cham.

Walter, S. L., Seibert, S. E., Goering, D., and O'Boyle, E. H. (2019). A tale of two sample sources: Do results from online panel data and conventional data converge? *Journal of Business and Psychology*, 34(4):425–452.

Walters, M. L., Koay, K. L., Syrdal, D. S., Dautenhahn, K., and Te Boekhorst, R. (2009). Preferences and perceptions of robot appearance and embodiment in human-robot interaction trials. *Procs of New Frontiers in Human-Robot Interaction*.

Weiss, A. and Bartneck, C. (2015). Meta analysis of the usage of the godspeed questionnaire series. In *2015 24th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*, pages 381–388. IEEE.

Wheeler, S. C. and Omair, A. (2016). Potential growth areas for implicit theories research. *Journal of consumer psychology*, 26(1):137–141.

Whelan, S. and Casey, D. (2021). Examining social robot acceptability for older adults and people with dementia. In Ortiz, J. H. and Vinjamuri, R. K., editors, *Collaborative and Humanoid Robots*, chapter 2. IntechOpen, Rijeka.

Williams, S. L. (2015). *Once a thief, always a thief? How time, implicit theories, and race affect moral judgments*. Unpublished doctoral thesis, University of Illinois at Urbana-Champaign.

Wullenkord, R. and Eyssel, F. (2020). The influence of robot number on robot group perception—a call for action. *ACM Transactions on Human-Robot Interaction (THRI)*, 9(4):1–14.

Wurthmann, K. (2013). Perceivers' implicit theories, violation type, and moral awareness. In *Academy of Management Proceedings*, volume 2013, page 13827. Academy of Management Briarcliff Manor, NY 10510.

Xia, Y. and LeTendre, G. (2020). Robots for future classrooms: A cross-cultural validation study of "negative attitudes toward robots scale" in the us context. *International Journal of Social Robotics*, pages 1–12.

Xu, K. (2019). First encounter with robot alpha: How individual differences interact with vocal and kinetic cues in users' social responses. *New Media & Society*, page 1461444819851479.

Yeager, D. S. and Dweck, C. S. (2020). What can be learned from growth mindset controversies? *American Psychologist*, 75(9):1269.

Yeager, D. S., Hanselman, P., Walton, G. M., Murray, J. S., Crosnoe, R., Muller, C., Tipton, E., Schneider, B., Hulleman, C. S., and Hinojosa, C. P. (2019). A national experiment reveals where a growth mindset improves achievement. *Nature*, pages 1–6.

Yeager, D. S., Miu, A. S., Powers, J., and Dweck, C. S. (2013). Implicit theories of personality and attributions of hostile intent: A meta-analysis, an experiment, and a longitudinal intervention. *Child development*, 84(5):1651–1667.

Yeager, D. S., Trzesniewski, K. H., Tirri, K., Nokelainen, P., and Dweck, C. S. (2011). Adolescents' implicit theories predict desire for vengeance after peer conflicts: Correlational and experimental evidence. *Developmental psychology*, 47(4):1090.

Yeager, D. S. and Walton, G. M. (2011). Social-psychological interventions in education: They're not magic. *Review of educational Research*, 81(2):267–301.

Yogeeswaran, K., Złotowski, J., Livingstone, M., Bartneck, C., Sumioka, H., and Ishiguro, H. (2016). The interactive effects of robot anthropomorphism and robot ability on perceived threat and support for robotics research. *Journal of Human-Robot Interaction*, 5(2):29–47.

Yorkston, E. A., Nunes, J. C., and Matta, S. (2010). The malleable brand: The role of implicit theories in evaluating brand extensions. *Journal of Marketing*, 74(1):80–93.

Zafari, S., Schwaninger, I., Hirschmanner, M., Schmidbauer, C., Weiss, A., and Koeszegi, S. T. (2019). "you are doing so great!"–the effect of a robot's interaction style on self-efficacy in hri. In *2019 28th IEEE International Conference on Robot and Human Interactive Communication (RO-MAN)*, pages 1–7. IEEE.

Zhao, S. (2006). Humanoid social robots as a medium of communication. *New Media & Society*, 8(3):401–419.

Zhao, Z. and McEwen, R. (2022). " let's read a book together" a long-term study on the usage of pre-school children with their home companion robot. In *Proceedings of the 2022 ACM/IEEE International Conference on Human-Robot Interaction*, pages 24–32.

Zhong, V. J., Mürset, N., Jäger, J., and Schmiedel, T. (2022). Exploring variables that affect robot likeability. In *Proceedings of the 2022 ACM/IEEE International Conference on Human-Robot Interaction*, page 1140–1145. IEEE Press.

Zimmerman, M., Bagchi, S., Marvel, J., and Nguyen, V. (2022). An analysis of metrics and methods in research from human-robot interaction conferences, 2015-2021. In *Proceedings of the 2022 ACM/IEEE International Conference on Human-Robot Interaction*, HRI '22, page 644–648. IEEE Press.

Złotowski, J., Yogeeswaran, K., and Bartneck, C. (2017). Can we control it? autonomous robots threaten human identity, uniqueness, safety, and resources. *International Journal of Human-Computer Studies*, 100:48–54.