



# Attitudes Toward Attributed Agency: Role of Perceived Control

Setareh Zafari<sup>1</sup> · Sabine T. Koeszegi<sup>1</sup>

Accepted: 23 June 2020 / Published online: 10 July 2020  
© The Author(s) 2020

## Abstract

Previous research suggests that the increased attribution of agency to robots may be linked to negative attitudes toward robots. If robots are truly expected to assume various roles in our social environment, it is necessary to further explore how increasing agency, for example through increasing levels of autonomy, affects attitudes toward them. This study investigates the role of perceived control as a moderator explaining attitudes toward attributed agency in a collaboration context. Austrian-based participants ( $N = 102$ ) watched a video of a robot collaborating with a person to assemble a mixer—the robot was presented as either agentic and capable of proactively collaborating with the human or non-agentic and only capable of following human commands. The results show that attributing high levels of agency to robots is associated with negative attitudes toward them when individuals perceive low control during the collaboration.

**Keywords** Attitudes toward robots · Attributed agency · Perceived control · Human–robot collaboration

## 1 Introduction

Collaboration between humans and robots is becoming more feasible thanks to advancements in robotics. In accordance with the vision of a cyber-society, autonomous robots are being used to assist with different activities in close contact with people in contexts ranging from workplaces (e.g. robots assembling automobiles in the manufacturing sector) to people's daily home lives (e.g. a service robot helping in the kitchen or feeding a patient in need of care). Although the technology is not yet sufficiently mature to be implemented widely, examples such as the ones above are becoming more common. This requires a better understanding of robots' impact on human beings. We are particularly concerned that some people might have difficulty accepting such a cooperative relationship with robots. One way to study the social acceptance of such robots is to focus on attitudes toward them [53,55]. Currently, there are negative attitudes in popular culture toward autonomous robots [20,34,69]. Robot abuse and antagonism toward robots can also be considered indicative of negative attitudes toward them. Given that the acceptance of robots in everyday life depends not only on technical but also on social and psychological aspects of human-robot interactions [35], it is necessary to first investi-

gate what causes these negative attitudes and then counteract mechanisms to improve them.

When interacting with novel entities such as autonomous robots, people tend to use existing social schemas to make sense of the situation. Perceiving a mind in robots is an example of this projection, or tendency to attribute human-like characteristics, motivations and intentions to non-human entities [64]. Mind perception has two dimensions: experience (the ability to feel emotions such as pleasure) and agency (the ability to act in ways such as self-control) [25,26]. Although primary work on uncanny feelings demonstrated that perceptions of experience of robot generate uncanny feelings [26], other recent studies (e.g. [3,69]) have shown that the robots' ability to act (i.e. perception of agency) also increases uncanny feelings. As Wallach and Allen state: “within the next few years we predict there will be a catastrophic incident brought by a computer system making a decision independent of human oversight” [62, p. 4]. Overall, these views suggest that robots' improved capability to autonomously act or make decisions would induce negative attitudes toward robots. Nevertheless, despite the negative press given to autonomous robots, research demonstrating a link between agency and negative attitudes is inconsistent. The current literature offers contradictory findings about the effects of agency attributed to robots. While several studies reported positive attitudes toward high agency [48,66], others found negative attitudes and outcomes [29,53]. As a result,

✉ Setareh Zafari  
setareh.zafari@tuwien.ac.at

<sup>1</sup> Institute of Management Science, TU Wien, Vienna, Austria

we have little understanding about the underlying mechanisms through which attributed agency to robots influences peoples' behavior and attitudes toward robots.

In this study, we address this gap by investigating the effect of perceived control. We expect that perceived control moderates the impact of attributed agency and may help to explain some of the heterogeneity in reactions and attitudes toward robots. Perceived control is relevant to understanding the consequences of the increased attribution of agency to robots for three reasons. First, the High-Level Expert Group has identified human oversight as a key requirement in their guidelines on minimizing the negative effects of AI systems [16]. However, few studies have investigated the role of human oversight and perception of control in collaboration with agentic robots in any systematic way. Second, studies on social robot acceptance reveal that characteristics of both robots and humans influence attitudes toward robots. For instance, physical embodiment, consistent substrate and a match between a robot's appearance and behavior foster users' acceptance of the robot [22,23,63]. Other research indicates that men and younger people tend to have more positive attitudes toward robots [27,35,59,60] than women and older people. While these demographic data may predict non-acceptance, they are not able to explain negative reactions and non-acceptance [53]. In this respect, other external factors such as perceived control could explain the existence of situations in which a high attribution of agency to robots does not imply negative attitudes and non-acceptance. Third, in the human-computer interactions (HCI) context, perceived control is defined as "the perception that one's behavior significantly alters outcomes by producing desired and preventing undesired events" [30, p. 4]. Given the diverse effects of control on human cognition, attitudes, and behavior, it should not be surprising that emerging technology undermining humans' control may be unpleasant. More specifically, such a relationship may be theoretically supported by the concept of reactance, the tendency to react negatively toward threats to one's behavioral freedom, whether in the form of eliminating or limiting it. According to reactance theory [8], people become aroused as a result of loss of control over the situation, which may lead them to feel aggressively toward the entity that is attempting to restrict their freedom. The present study examines whether attitudes toward agentic robots depend on the control humans feel during the collaboration. Thus, this study's contributions to the human-robot interaction (HRI) literature are twofold. Firstly, a relatively new theoretical conceptualization of control, namely perceived control as a situational appraisal, is used. Secondly, in this study, we test for moderation in the relationship between attributed agency and attitudes toward robots.

We specifically focus on collaboration with autonomous robots as an example of interacting with agentic robots. Agency is a characteristic that refers to perceived auton-

omy in robots' behavior [45]. In a collaborative context, a robot that demonstrates proactive behaviors seems to have more agency compared to situations in which the robot only responds to orders (i.e. reactive behavior). We understand agency as the capacity to perform a goal-oriented task to an extent autonomously on the environment [68]. Rose and Turex describe machine agency as the extent to which machines are perceived by humans as having autonomy [45]. Accordingly, there is an undeniable relationship between autonomy and agency: as the autonomy level of a robot increases, greater agency is attributed to the robot. Thus, collaboration with a robot with a higher level of autonomy provokes a higher attribution of agency to the robot.

## 2 Background and Related Work

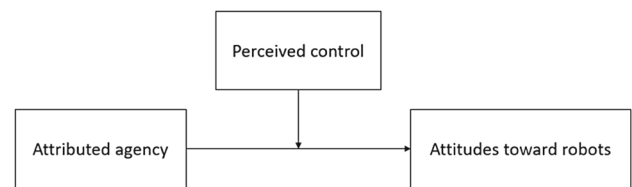
Human-robot collaboration refers to a collaborative accomplishment by humans and robots in completing tasks with a focus on coordinating close, seamless joint activities between humans and robots [1]. The participants' mutual engagement enables problem solving that cannot be achieved without the two sides' direct coordination and interaction. Furthermore, the focus of Human-robot collaboration is not to replace a human by a robot but to complement each other by contributing their strengths toward the mutual goal that is stated by the human agent [33]. Given the increasing interest in understanding dynamics beyond mere implementation, in this study, we discuss the contextual conditions under which collaboration with agentic robots produces negative attitudes toward robots.

Prior studies of robots in work contexts provide some preliminary evidence for the effect of increased agency on acceptance and attitudes toward robots. Stafford reported that people were more likely to use robots when they perceived robots' minds as having less agency [53]. Heerink et al. reported more anxiety toward a more adaptive robot than a less adaptive one [29]. However, the effects of collaboration with agentic robot are complex and may not lead only to negative outcomes. Different contextual factors may influence the relationship between agency attributed to robots and attitudes toward them. For instance, Wiese et al. show that people were more willing to engage in joint attention with a robot when they treat it as a system with intentionality [66]. A study by Rau et al. found marginal effect of positive effect of level of autonomy the robot's influence on human decision when the robot was an in-group compared to an out-group member [44]. Liu et al. found that participants preferred to work with a robot that adapted to their actions over one that did not [37]. In another study conducted by Shermerhorn and Scheutz, people attribute greater cooperativeness to a robot in autonomous mode and accept dynamic autonomy when the robot makes autonomous decisions in the interest

of team goals, even going so far as to ignore instances of disobedience [48]. While these studies have provided important information about how people interact with autonomous robots, the existing evidence does not conclusively determine how an increased attribution of agency of may affect peoples' attitudes toward the robot. Given that contextual conditions can shape how individuals develop attitudes toward robots, perceived control may be a particularly important dimension of cognition to investigate.

The contextual variable investigated in this paper is perceived control. Our notion of perceived control is different from other interpretations of control, such as self-efficacy in HRI [43] or perceived behavioral control in the theory of planned behavior [2], which reflects users' perceived ease of performing a behavior and treats the perceived difficulty and perceived control as the same construct. Researchers argue that these two have different antecedents and should be considered as separate case. Perceived control is determined by a set of underlying control beliefs that mostly capture external factors (situational influence), while perceived difficulty is determined by internal factors (ability and skills) and therefore focuses on self-efficacy [13]. Perceived control is defined as “appraisal of the extent to which other people or events will interfere with the performance of the behavior” [57, p. 202]. It is an individual's interpretation and belief about how much control is available [52]. Moreover, research on the illusion of control has emphasized the importance of perceived control over locus of control [36]. While locus of control refers to individuals' general beliefs about the main causes of events in their lives (i.e. external or internal locus of control) [47], perceived control refers to a more situational perceived ability to affect the outcome of a course of action [41,42]. Few studies have examined the concept of perceived control in the HRI contexts [12,32,38]. While these results indicate that people prefer to be in control, it is unclear what this means for human–robot collaboration. In this context, humans can undertake action or decisions indirectly through a robot, which would cause the person to no longer consider themselves as in control of the action. Feeling that one is in control may improve the sense of agency, when the achieved outcome of an action conforms to one's intention [42]. According to Pacherie, sense of control is one of the main contributor to sense of agency (the sense that an individual is the author of an action) and comprises three more basic experiences: sense of motor control, sense of situational control and sense of rational control [42]. While situational control is perceptual and represents “control of the action with regard to the situation as currently perceived” [42, p. 4], rational control exists at a higher level of abstraction and represents a more global consistency that is not affected by a single event or experience.

We predict that increases in robot's agency are therefore likely to have a negative affect on attitudes toward robots if



**Fig. 1** Our conceptual model in which the perceived control moderates the relationship between attributed agency and attitudes toward robots

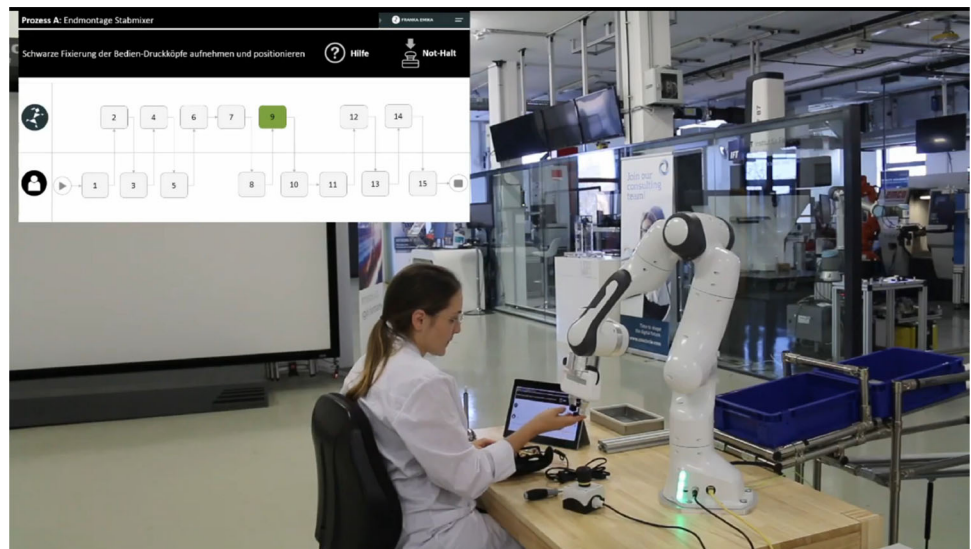
humans feel less in control. This proposition receives support from [40], who argues that an increase in the level of autonomy is related to negative emotions in users because it leads them to experience a lack of control. Studies have shown negative arousal in social experiences and environmental conditions that can potentially diminish an individual's perceived control [69]. Moreover, people who perceive high control are more confident about their performance. People who perceive higher control may feel more comfortable collaborating with a robot because they do not see it as undermining their values or contribution to the task [5]. As suggested by Hinds, when control appears to be in the hand of another entity, people experience negative feelings, which could lead to lower acceptance or use of that system [30]. Moreover, in reactance theory [8], individuals react negatively to threats to their control. Therefore, conditions that reduce the perception of control may lead to motivational and cognitive deficits [54] and consequently negative attitudes. Given that the perception of a robot's agency increases as its autonomy rises [56], it is certainly possible that the positive relationship between a robot's attributed agency and negative attitudes is among individuals who feel low control over the task performance. Thus, we expect those who were exposed to agentic robots and who believed that they had little control reported greater negative attitudes toward robots relative to individuals exposed to non-agentic (see Fig. 1).

## 3 Methodology

### 3.1 Participants

102 participants were recruited through snowball sampling for this study in order to expand the sample size and scope. 9 participants were excluded for failing the manipulation check of attributing the correct amount of agency to the robot (i.e. the video failed to elicit the targeted level of the robot's perceived autonomy in the participant). One outlier was also detected and removed. Thus, the final sample consisted of a total of 92 participants (46% women and 54% men), ranging in age from 20 to 63 years old ( $M = 32.93$ ,  $SD = 9.86$ ). The majority of respondents were highly educated (i.e. 38% held a master's degree). Half of the participants (50%) indicated

**Fig. 2** Example screenshot of the video vignette that presents a collaboration between a person and a robot



having previous interaction experience with a robot. The data was collected between 5 and 23 February 2020.

### 3.2 Manipulation

To assess people's attitudes toward the attributed agency of robots, we simulated a collaboration process regarding assembling a product in a factory using a vignette study. Video vignette was developed in which participants were asked to imagine they are a member of a human–robot manufacturing team; the manufacturer had recently acquired a new robot to work alongside people to improve their productivity. The goal, building a mixer, was set by the manufacturer and assigned to this team. This goal can only be achieved if the two agents collaborate with each other. Given the product's complexity, some of the tasks will be done by the robot, and other tasks will be done by the human participant. Figure 2 shows a screenshot of the video vignette that participants were asked to watch.

We manipulated level of agency attributed to the robot as a between-subjects factor by varying the level of autonomy in the robot's behavior. Two experimental conditions (i.e. HIGH agency and LOW agency) were implemented in the robot system to be evaluated. In the LOW agency condition<sup>1</sup>, the robot reactively participated in decision-making and completion of the product, i.e. the employee made the decisions about the sequence of actions required to assemble the mixer, and the robot required human commands to complete its assigned actions. In the case of a problem or ambiguity, the person in the video had the option of asking for help from the robot, which the robot could only make suggestions about what to do. The robot communicated with the

human agent through text on the tablet screen. In the HIGH agency condition<sup>2</sup>, it was the robot who decided the sequence of actions, told the employee which tasks to carry out and the user only had the right to veto the robot's decisions. In the case of a problem or work stoppage, the robot proactively shows the person how to do the task without being asked. The main difference between the conditions is that the robot's behaviour in the former is relatively deterministic, while its behavior in the latter is unpredictable and actions are not suggested but are used in imperative manner to reflect the higher level of autonomy. A pilot test involving 40 students at TU Vienna was carried out to verify the feasibility and validity of the vignette. In addition to question items mention below in the measures section, participants were asked to describe in their own words what they have seen in the video. As a consequence from the pilot test, we added further information about the robot arm at the beginning of the video and questionnaire as some of the pilot participants had difficulties recognizing the interactive tablet belongs to the robot arm.

### 3.3 Procedure

Participants were randomly assigned to watch one of two videos uploaded as private videos on Youtube. After reading an information sheet that outlined the purpose of the study and providing consent to participate, participants were randomly assigned to watch a video of a person collaborating with either an agentic (HIGH agency) or non-agentic (LOW agency) robot in an assembly task. We asked people to respond from the vignette character's perspective as if they were that person in that situation, rather than on the basis of their own lives. This should

<sup>1</sup> <https://www.youtube.com/embed/fg4MG9r5Pvs>.

<sup>2</sup> <https://www.youtube.com/embed/1ykbuSZ7BxY>.

help to reduce the effects of socially desirable response patterns [31]. After watching the video, participants completed a post-video questionnaire. The survey instruments were all provided in German. The questionnaire were constructed using a double translation procedure conducted by two different researchers fluent in both German and English.

### 3.4 Measures

For the manipulation check, participants were asked to indicate the level of the robot's perceived autonomy on a 5-point Likert scale ranging from 1 (= low) to 5 (= high).

Attitudes toward robots were measured with 14 items from the Negative attitudes toward Robots Scale (NARS) [39] and 11 items from the Robot Attitudes Scale (RAS) [9]. These scales have been applied successfully to measure the psychological reactions to (human like and non-human like) robots in a general sense but also after interacting with a specific robot (e.g. [50,65]). The NARS contains 3 subscales that measure negative attitudes towards (1) situations of interaction with robots (NARS-S1) (a sample item is "I would feel uneasy if I was given a job where I had to use robots"), (2) the social influence of robots (NARS-S2) (a sample item is "I feel that in the future society will be dominated by robots"), and (3) emotions in interaction with robots (NARS-S3) (a sample item is "I would feel relaxed talking with robots"). All items were rated on a 5-point Likert-type scale ranging from 1 (= strongly disagree) to 5 (= strongly agree). A higher score on this scale indicates a more negative attitude. The reported Cronbach's alpha for the NARS-S1, NARS-S2, and NARS-S3 were 0.72, 0.70, and 0.72, respectively. The official German adaptation of the questionnaire was used [7]. The RAS was used to rate what the participants thought of robots on scales from 1–8. Items included the adjectives friendly, useful, trustworthy, strong, interesting, advanced, easy to use, reliable, safe, simple, and helpful. Higher scores on this scale are associated with less favourable attitudes toward robots. Cronbach's alpha was 0.82.

To measure perceived control, we used 2 items from [30]. A sample item is "I felt that I was in control". These items were highly correlated ( $r=0.71$ ,  $p < 0.01$ ).

Participants reported basic socio-demographic information such as age, gender (0 = female, 1 = male), level of education and previous experience with robots (0 = no, 1 = yes). A personality trait which has been shown to be related to perceived control is the desire to control [10,24]. Individual differences in the level of motivation to control the events in one's life were measured with 20 items from the Desirability of Control Scale (DC) [11]. Cronbach's alpha was 0.81. All items were rated on a 7-point Likert-type scale ranging from 1 (= strongly disagree) to 7 (= strongly agree).

## 4 Results

An analysis of standard residuals was carried out to identify any outliers, which indicated that participants 81 needed to be removed. The histogram of standardised residuals indicated that the data contained approximately normally distributed errors, as did the normal P-P plot of standardised residuals. Furthermore, the assumption of homogeneity of variance was checked using Levene's test. Multicollinearity between the independent variables was examined using the VIF coefficients. The results showed that multicollinearity was not a concern.

Table 1 displays the means, standard deviations, and correlations of our study variables. There was no evidence of a direct relationship between attributed agency and attitudes toward robots, at least expressed in terms of simple association. A weak significant correlation between perceived control and RAS was found ( $r = -0.28$ ,  $p < 0.01$ ). As perceived control increases, negative rating of the robot decreases. We also found a moderate significant correlation between the extent to which they attributed agency to the robot and how much they perceived control ( $r = -0.42$ ,  $p < 0.01$ ), meaning that the higher the level of agency resulted in lower perceived control. Significant correlations between gender and RAS ( $r = 0.21$ ,  $p < 0.05$ ), gender and NARS-S1 ( $r = -0.34$ ,  $p < 0.01$ ), gender and NARS-S2 ( $r = -0.30$ ,  $p < 0.01$ ), previous experience with robots and NARS-S1 ( $r = -0.36$ ,  $p < 0.01$ ), previous experience with robots and NARS-S1 ( $r = -0.21$ ,  $p < 0.05$ ), perceived control and age ( $r = 0.24$ ,  $p < 0.05$ ), perceived control and gender ( $r = -0.23$ ,  $p < 0.05$ ), and were also found. Age, education level and desirability to control were not associated with attitudes toward robots.

An independent-samples t-test was conducted to compare attitudes toward robots in the LOW and HIGH agency conditions. Lower RAS scores were reported for LOW agency ( $M = 2.88$ ,  $SD = 1.01$ ) than HIGH agency ( $M = 3.37$ ,  $SD = 1.08$ ) condition;  $t(90) = -2.25$ ,  $p = 0.03$ . However, there were no significant differences in the NARS-S1 (HIGH agency:  $M = 2.10$ ,  $SD = 0.86$ ; LOW agency:  $M = 1.86$ ,  $SD = 0.62$ ;  $t(90) = -1.51$ , n.s.), NARS-S2 (HIGH agency:  $M = 2.60$ ,  $SD = 0.95$ ; LOW agency:  $M = 2.65$ ,  $SD = 0.67$ ;  $t(90) = 0.34$ , n.s.) and NARS-S3 scores (HIGH agency:  $M = 3.20$ ,  $SD = 0.89$ ; LOW agency:  $M = 3.53$ ,  $SD = 0.95$ ;  $t(90) = 1.68$ , n.s.).

To investigate whether perceived control moderates the relation between attributed agency and attitudes toward robots, the SPSS script (Model 1) by Preacher and Hayes [28] was used. The results were tested using 1000 bootstrapped samples and 95 percent confidence intervals. Age, gender, level of education, previous experience with robots and desirability of control were entered as covariates.

As predicted, perceived control moderated the relationship between attributed agency and NARS-S1 (*coeff.* =

**Table 1** Means, standard deviations and correlations for the study variables

	M	SD	1	2	3	4	5	6	7	8	9	10
1. NARS-S1	1.96	0.74										
2. NARS-S2	2.63	0.79	0.66**									
3. NARS-S3	3.39	0.94	0.16	0.28**								
4. RAS	3.09	1.06	0.28**	0.15	−0.10							
5. Attributed agency	2.63	1.13	0.20	0.20	−0.01	0.19						
6. Perceived control	4.49	1.91	−0.17	−0.05	0.13	−0.28**	−0.42**					
7. Age	32.93	9.86	0.06	0.02	−0.11	−0.03	−0.03	−0.24*				
8. Gender	0.54	0.50	−0.34**	−0.30**	−0.08	0.21*	−0.01	−0.23*	−0.02			
9. Education level	6.87	1.48	−0.13	−0.02	−0.10	−0.06	−0.01	−0.01	0.21*	0.01		
10. Pre. Exp. with robots	0.50	0.50	−0.36**	−0.22*	−0.10	0.01	−0.08	−0.05	−0.07	0.35**	0.07	
11. Desirability of control	4.81	0.74	−0.19	−0.15	−0.01	0.03	−0.01	−0.01	0.20	0.14	0.13	0.30**

NARS-S1: negative attitudes towards situations of interaction with robots

NARS-S2: negative attitudes towards social influence of robots

NARS-S3: negative attitudes towards emotions in interaction with robots

\* $p < 0.05$ , \*\* $p < 0.01$ **Table 2** Moderation analysis results for attributed agency as independent variable

	NARS-S1			NARS-S2			NARS-S3			RAS		
	Est.	SE	t	Est.	SE	t	Est.	SE	t	Est.	SE	t
Intercept	1.67	0.76	2.19*	1.76	0.91	1.93	2.10	1.15	1.83	2.04	1.24	1.64
Age	0.00	0.01	0.47	0.00	0.01	0.37	−0.01	0.01	−0.63	−0.01	0.01	−0.79
Gender	−0.41	0.14	−2.86**	−0.39	0.17	−2.30*	−0.03	0.22	0.42	0.42	0.23	1.81
Education level	−0.03	0.05	−0.70	0.02	0.05	0.28	−0.03	0.07	−0.44	−0.01	0.07	−0.17
Prev. exp. with robots	−0.31	0.15	−2.13*	−0.10	0.17	−0.58	−0.16	0.22	−0.71	−0.18	0.24	−0.77
Desirability of control	−0.08	0.10	−0.83	−0.11	0.11	−0.95	0.06	0.14	0.44	0.07	0.16	0.47
A. Agency	0.53	0.15	3.41**	0.53	0.18	2.88**	0.44	0.23	1.92	0.53	0.25	2.11*
P. Control	0.21	0.09	2.18*	0.23	0.11	2.05*	0.31	0.14	2.19*	0.15	0.15	0.96
A. Agency* P. Control	−0.10	0.03	−3.34**	−0.09	0.04	−2.43*	−0.09	0.05	−2.01*	−0.10	0.05	−1.97*
R <sup>2</sup>	0.35			0.20			0.08			0.17		
F-value	5.60***			2.64*			0.97			1.03*		

A. Agency: attributed agency; P. Control: perceived control

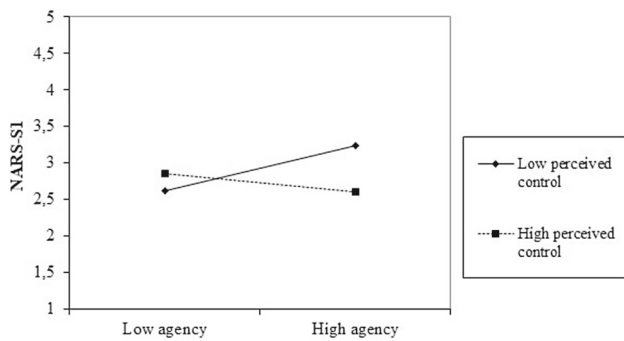
\* $p < 0.05$ , \*\* $p < 0.01$ , \*\*\* $p < 0.001$ 

−0.10,  $p < 0.01$ , see Table 2). As Fig. 3 shows, when perceived control was low, there was a positive relationship between agency and NARS-S1. That is, for the individuals who perceived lower control, the more they attributed agency to the robot, the more negative they were regarding situations of interaction with robots.

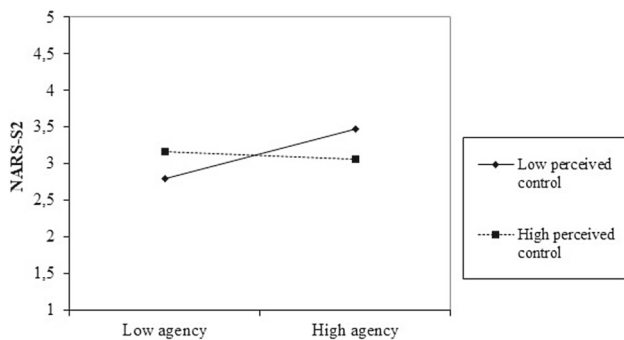
The moderating effect of perceived control on the relationship between attributed agency and NARS-S2 was also significant ( $coeff. = -0.09$ ,  $p < 0.05$ , see Table 2). As shown in Fig. 4, when perceived control was low, the relationship between agency and NARS-S2 was positive. That is, for the individuals who perceived lower control, the more they attributed agency to the robot, the more negative they were regarding social influence of robots.

The moderating effect of perceived control on attributed agency and NARS-S3 ( $coeff. = -0.09$ ,  $p < 0.05$ , see Table 2) was significant. However, the model was not significant ( $F(8,83) = 0.97$ , n.s.).

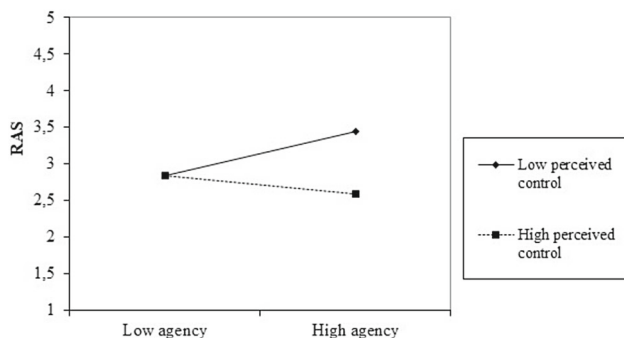
Perceived control also moderated the relationship between attributed agency and RAS ( $coeff. = -0.10$ ,  $p < 0.05$ , see Table 2). Figure 5 shows that when perceived control was low, there was a positive relationship between agency and RAS, indicating that greater attributed agency was associated with less favorable attitudes toward robots among those who felt they have little control over the task.



**Fig. 3** Interaction of attributed agency and perceived control on NARS-S1



**Fig. 4** Interaction of attributed agency and perceived control on NARS-S2



**Fig. 5** Interaction of attributed agency and perceived control on RAS

## 5 Discussion

In this study, we focused on how agency attributed to robots is related to attitudes toward them by investigating the intervening role of perceived control in a collaboration process. In particular, we examined whether perceived control moderates the relationship between attributed agency and attitudes toward robots.

The results of our study confirmed our hypothesis for RAS and two subscales of NARS (i.e. NARS-S1 and NARS-S2) that the relationship between attributed agency and attitudes is contingent on perceived control. For NARS-S3, despite a significant interaction term between perceived control and

attributed agency, the model was not significant. We found that when perceived control is low, there is a positive relationship between attributed agency and negative attitudes toward interaction situations with robots as well as the social influence of robots. This indicates that those who attributed relatively higher agency to the robot did report relatively less favorable attitudes toward robots compared to those attributed lower agency, but this was true only among those with relatively low perceived control. These findings can be explained by the stress literature [19,21], which indicates that lack of control undermines individuals' ability to cope with a stressful situation and even the mere perception of control (i.e. without control actually being available) can also reduce stress. Consistent with reactance theory [8], our results show that a lack of control results in people having negative attitudes toward robots.

We also observed some demographic differences regarding attitudes towards robots. Consistent with the literature [17,59], this study found that female respondents and individuals with no previous interaction with robot were more likely to report negative attitudes toward robots. Moreover, this study supports previous observations (e.g. [51,59]) finding no significant correlation between age and attitudes toward robots.

In accordance with prior studies noting the importance of individual subjective attributions of agency [18,53,69] in explaining attitudes toward robots, the aim of our study was to further investigate psychological factors, such as perceived control, involved in the acceptance of robots. Numerous efforts are focused on including humans in the decision-making loop to improve the quality of task plans and schedules for autonomous systems [4,15]. To our knowledge, the current study was the first to empirically investigate the role of perceived control and human oversight in a human-robot collaboration context. Our results yield important insights about how human cognition is developing alongside the robots we are creating, helping us to understand factors that facilitate or hinder their social acceptance. Given that individuals' preference for control has been found to increase over time as they get used to the robot [32], it is necessary to provide cognitive and behavioral resources for individuals who work alongside robots. Thompson identifies factors that can enhance perceptions of control, such as assessing skill acquisition, costs and benefits, the accuracy of self-efficacy expectations, etc. [58]. Accordingly, we can infer that if the context encourages individuals to perceive more control in their collaboration with robots, the relationship between attributed agency and positive attitudes should be positive. In this study, we found that female and younger respondents are more likely to experience higher perceived control. A further study with a focus on other factors affecting the perception of control in human-robot collaboration is therefore suggested. Furthermore, this study showed that perception of control

could be induced more strongly when actions are selected by the individual as opposed to instructed by a robot. A further potential direction for future studies could be to determine how the valence of an action can affect the perceived control. Experimental research on neurobiology have shown that unpleasant outcomes lead to lower feeling of control and consequently lower sense of agency in people compared to positive outcome [6,67]. These mechanisms need to be further investigated in HRI context, as recent studies [14,46] found that the social presence of robot reduces the sense of agency over self-generated actions. Since enhancement of human agency is necessary to protect human rights [61], further studies critically investigating the effect of robots on our sense of agency and perception of control would be worthwhile.

This study is subject to several limitations. First, there is a need to investigate interaction contexts where participants are interactants and not observers. The social interaction literature suggests that social cognition may be different when the person is in the role of an interactant rather than an observer of interactions [49]. Our results therefore need to be interpreted with caution. Future research could seek to verify our current findings in real human-robot interaction scenarios to gain a better understanding of the buffering role of perceived control. Second, the analysis was based solely on Austrian data, and the results cannot be generalised directly to other countries. The literature suggests that responses to robot-related attitudes questions reflect respondents' individual experiences, and perspectives and may be susceptible to cultural differences [39,55] and a country's technological orientation [59]. However, it is unlikely to suppose that control perceptions would be dependent on culture or country in the same manner. In addition, this study focused on perceived control and tested how it acts as a boundary condition for the positive relationship between attributed agency and attitudes toward robots. As individuals' reaction to perceived control rests on their motive to control [10], future research might integrate desire for control in the model and investigation the effects of match and mismatch between perceived control and desire for control.

Consistent with other studies [35,44], our research further highlights the importance of including both technical and social aspects in designing robots. The findings may help illuminate the process by which agency attributed to a robot is linked to the development of negative attitudes towards them. When individuals feel in control and believe that they themselves determine the task outcome and not others' actions or external factors, they tend to feel more comfortable in collaborating with robots. Thus, beyond understanding whether the attribution of agency leads to negative attitudes toward robots, future studies should consider the nature of perceived control in the collaborative context, which explains when

agency attributed to the robots is associated with positive attitudes.

## 6 Conclusion

The purpose of the current study was to examine conditions under which a robot's attributed agency is associated with negative attitudes toward robots by addressing the role of perceived control. The results indicate that increases in attributed agency are associated with negative attitudes toward robots when individuals feel lack of control in a collaboration context with robots. The findings suggest that perceived control can mitigate negative attitudes and foster social relationship between humans and robots.

**Acknowledgements** Open access funding provided by TU Wien (TUV). The authors wish to thank the study participants for their collaboration and Christina Schmidbauer for her support in setting up the study.

## Compliance with Ethical Standards

**Conflict of interest** The authors declare that they have no conflict of interest.

**Research Involving Human Participants** The authors have sought advice from the Research Ethics Coordinator and the Data Protection Experts at TU Wien as there is at a moment - no institutional research ethics committee to approach for approval.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

1. Ajoudani A, Zanchettin AM, Ivaldi S, Albu-Schäffer A, Kosuge K, Khatib O (2018) Progress and prospects of the human-robot collaboration. *Autonomous Robots* 42(5):957–975
2. Ajzen I (1985) From intentions to actions: a theory of planned behavior. In: *Action control*, pp 11–39. Springer, Berlin
3. Appel M, Weber S, Krause S, Mara M (2016) On the eeriness of service robots with emotional capabilities. In: *The Eleventh ACM/IEEE international conference on human robot interaction*, pp 411–412. IEEE Press
4. Ardissono L, Petrone G, Torta G, Segnan M (2012) Mixed-initiative scheduling of tasks in user collaboration. In: *WEBIST*, pp 342–351

5. Bandura A (1986) Social foundations of thought and action. Englewood Cliffs, New Jersey
6. Barlas Z, Kopp S (2018) Action choice and outcome congruency independently affect intentional binding and feeling of control judgments. *Front Human Neurosci* 12:137
7. Bartneck C, Nomura T, Kanda T, Suzuki T, Kennsuke K (2005) A cross-cultural study on attitudes towards robots. In: *HCI international*
8. Brehm JW (1966) A theory of psychological reactance. Academic Press, Cambridge
9. Broadbent E, Tamagawa R, Kerse N, Knock B, Patience A, MacDonald B (2009) Retirement home staff and residents' preferences for healthcare robots. In: *RO-MAN 2009-The 18th IEEE international symposium on robot and human interactive communication*, pp. 645–650. IEEE
10. Burger JM (1992) Desire for control: personality, social, and clinical perspectives. Springer, Berlin
11. Burger JM, Cooper HM (1979) The desirability of control. *Motivation Emotion* 3(4):381–393
12. Chanseau A, Dautenhahn K, Koay KL, Salem M (2016) Who is in charge? sense of control and robot anxiety in human-robot interaction. In: *2016 25th IEEE international symposium on robot and human interactive communication (RO-MAN)*, pp. 743–748. IEEE
13. Cheung SF, Chan DK (2000) The role of perceived behavioral control in predicting human behavior: a meta-analytic review of studies on the theory of planned behavior. Unpublished manuscript, Chinese University of Hong Kong
14. Ciardo F, De Tommaso D, Beyer F, Wykowska A (2018) Reduced sense of agency in human-robot interaction. In: *International conference on social robotics*, pp. 441–450. Springer, Berlin
15. Clare AS, Cummings ML, How JP, Whitten AK, Toupet O (2012) Operator object function guidance for a real-time unmanned vehicle scheduling algorithm. *J Aerospace Comput Inform Commun* 9(4):161–173
16. Commission E (2020) White paper on artificial intelligence—a European approach to excellence and trust. White paper, European Commission
17. De Graaf MM, Allouch SB (2013) Exploring influencing variables for the acceptance of social robots. *Robotics Autonomous Syst* 61(12):1476–1486
18. Echterhoff G, Böhner G, Siebler F (2006) Social robotics und mensch-maschine-interaktion. *Z Sozialpsychologie* 37(4):219–231
19. Endler NS, Speer RL, Johnson JM, Flett GL (2000) Controllability, coping, efficacy, and distress. *Eur J Pers* 14(3):245–264
20. Flandorfer P (2012) Population ageing and socially assistive robots for elderly persons: the importance of sociodemographic factors for user acceptance. *Int J Popul Res*. <https://doi.org/10.1155/2012/829835>
21. Folkman S (1988) Personal control and stress and coping processes: a theoretical analysis. *Kango Kenkyu Jpn J Nurs Res* 21(3):243
22. Fong T, Nourbakhsh I, Dautenhahn K (2003) A survey of socially interactive robots. *Robotics Autonomous Syst* 42(3–4):143–166
23. Goetz J, Kiesler S, 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.*, pp. 55–60. IEEE
24. Gombolay MC, Gutierrez RA, Clarke SG, Sturla GF, Shah JA (2015) Decision-making authority, team efficiency and human worker satisfaction in mixed human-robot teams. *Autonomous Robots* 39(3):293–312
25. Gray HM, Gray K, Wegner DM (2007) Dimensions of mind perception. *Science* 315(5812):619–619
26. Gray K, Wegner DM (2012) Feeling robots and human zombies: mind perception and the uncanny valley. *Cognition* 125(1):125–130
27. Haring KS, Mougenot C, Ono F, Watanabe K (2014) Cultural differences in perception and attitude towards robots. *Int J Affect Eng* 13(3):149–157
28. Hayes A (2013) Introduction to mediation, moderation, and conditional process analysis: a regression-based approach. 2013. Guilford, New York
29. Heerink M, Kröse B, Wielinga B, Evers V (2008) Enjoyment intention to use and actual use of a conversational robot by elderly people. In: *Proceedings of the 3rd ACM/IEEE international conference on Human robot interaction*, pp 113–120. ACM
30. Hinds PJ (1998) User control and its many facets: a study of perceived control in human-computer interaction. Hewlett Packard Laboratories, California
31. Hughes R, Huby M (2012) The construction and interpretation of vignettes in social research. *Soc Work Soc Sci Rev* 11(1):36–51
32. Koay KL, Syrdal DS, Walters ML, Dautenhahn K (2007) Living with robots: investigating the habituation effect in participants' preferences during a longitudinal human-robot interaction study. In: *RO-MAN 2007-the 16th IEEE international symposium on robot and human interactive communication*, pp. 564–569. IEEE
33. Kolbeinsson A, Lagerstedt E, Lindblom J (2019) Foundation for a classification of collaboration levels for human-robot cooperation in manufacturing. *Prod Manuf Res* 7(1):448–471
34. Konok V, Korcsok B, Miklósi Á, Gácsi M (2018) Should we love robots?-the most liked qualities of companion dogs and how they can be implemented in social robots. *Comput Hum Behav* 80:132–142
35. Kuhnert B, Ragni M, Lindner F (2017) The gap between human's attitude towards robots in general and human's expectation of an ideal everyday life robot. In: *2017 26th IEEE international symposium on robot and human interactive communication (RO-MAN)*, pp. 1102–1107. IEEE
36. Langer EJ (1975) The illusion of control. *J Pers Soc Psychol* 32(2):311
37. Liu C, Hamrick JB, Fisac JF, Dragan AD, Hedrick JK, Sastry SS, Griffiths TL (2018) Goal inference improves objective and perceived performance in human-robot collaboration. *arXiv preprint arXiv:1802.01780*
38. Meerbeek BW, Saerbeck M, Bartneck C (2009) Iterative design process for robots with personality. *Society for the study of artificial intelligence and the simulation of behaviour (SSAISB)* pp. 94–101
39. Nomura T, Kanda T, Suzuki T (2006) Experimental investigation into influence of negative attitudes toward robots on human-robot interaction. *Ai Soc* 20(2):138–150
40. Norman DA (1994) How might people interact with agents. *Commun ACM* 37(7):68–71
41. Pacheco NA, Lunardo R, Santos CPd (2013) A perceived-control based model to understanding the effects of co-production on satisfaction. *BAR Braz Adm Rev* 10(2):219–238
42. Pacherie E (2007) The sense of control and the sense of agency. *Psyche* 13(1):1–30
43. Pütten ARVD, Bock N (2018) Development and validation of the self-efficacy in human-robot-interaction scale (se-hri). *ACM Trans Human-Robot Inter (THRI)* 7(3):1–30
44. Rau PLP, Li Y, Liu J (2013) Effects of a social robot's autonomy and group orientation on human decision-making. *Adv Human Comput Interaction*. <https://doi.org/10.1155/2013/263721>
45. Rose J, Truex D (2000) Machine agency as perceived autonomy: an action perspective. In: *Organizational and social perspectives on information technology*, pp. 371–388. Springer, Berlin
46. Roselli C, Ciardo F, Wykowska A (2019) Robots improve judgments on self-generated actions: an intentional binding study. In: *International conference on social robotics*, pp. 88–97. Springer

47. Rotter JB (1966) Generalized expectancies for internal versus external control of reinforcement. *Psychol Monogr General Appl* 80(1):1
48. Schermerhorn P, Scheutz M (2009) Dynamic robot autonomy: Investigating the effects of robot decision-making in a human-robot team task. In: *Proceedings of the 2009 international conference on multimodal interfaces*, pp. 63–70. ACM
49. Schilbach L, Timmermans B, Reddy V, Costall A, Bente G, Schlicht T, Vogeley K (2013) Toward a second-person neuroscience 1. *Behav Brain Sci* 36(4):393–414
50. Schneider S, Riether N, Berger I, Kummert F (2014) How socially assistive robots supporting on cognitive tasks perform. In: *Proceedings of the 50th anniversary convention of the AISB*, p. 35
51. Sinnema L, Alimardani M (2019) The attitude of elderly and young adults towards a humanoid robot as a facilitator for social interaction. In: *International conference on social robotics*, pp. 24–33. Springer
52. Skinner EA (1996) A guide to constructs of control. *J Pers Soc Psychol* 71(3):549
53. Stafford RQ, MacDonald BA, Jayawardena C, Wegner DM, Broadbent E (2014) Does the robot have a mind? mind perception and attitudes towards robots predict use of an eldercare robot. *Int J Social Robot* 6(1):17–32
54. Stanton JM, Barnes-Farrell JL (1996) Effects of electronic performance monitoring on personal control, task satisfaction, and task performance. *J Appl Psychol* 81(6):738
55. Syrdal DS, Dautenhahn K, Koay KL, Walters ML (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*
56. Takayama L (2012) Perspectives on agency interacting with and through personal robots. In: *Human-Computer interaction: the agency perspective*, pp. 195–214. Springer
57. Terry DJ, O’Leary JE (1995) The theory of planned behaviour: the effects of perceived behavioural control and self-efficacy. *Br J Soc Psychol* 34(2):199–220
58. Thompson SC (1991) *Handbook of social and clinical psychology: the health perspective* (Pergamon general psychology series), chap. Intervening to enhance perceptions of control, pp. 607–623. Pergamon Press, Oxford
59. Turja T, Oksanen A (2019) Robot acceptance at work: a multilevel analysis based on 27 EU countries. *Int J Soc Robot* 11(4):679–689
60. Venkatesh V, Morris MG, Davis GB, Davis FD (2003) User acceptance of information technology: toward a unified view. *MIS quarterly* pp. 425–478
61. Wagner B (2019) Liable, but not in control? ensuring meaningful human agency in automated decision-making systems. *Policy Internet* 11(1):104–122
62. Wallach W, Allen C (2008) *Moral machines: teaching robots right from wrong*. Oxford University Press, Oxford
63. Wang B, Rau PLP (2019) Influence of embodiment and substrate of social robots on users’ decision-making and attitude. *Int J Social Robot* 11(3):411–421
64. Waytz A, Gray K, Epley N, Wegner DM (2010) Causes and consequences of mind perception. *Trends Cognitive Sci* 14(8):383–388
65. Weiss A, Bernhaupt R, Tscheligi M, Yoshida E (2009) Addressing user experience and societal impact in a user study with a humanoid robot. In: *AISB2009: Proceedings of the symposium on new frontiers in human-robot interaction* (Edinburgh, 8–9 April 2009), SSAISB, pp. 150–157. Citeseer
66. Wiese E, Wykowska A, Zwickel J, Müller HJ (2012) I see what you mean: how attentional selection is shaped by ascribing intentions to others. *PLoS ONE* 7(9):e45391
67. Yoshie M, Haggard P (2013) Negative emotional outcomes attenuate sense of agency over voluntary actions. *Curr Biol* 23(20):2028–2032
68. Zafari S, Koeszegi ST (2018) Machine agency in socio-technical systems: a typology of autonomous artificial agents. In: *2018 IEEE Workshop on advanced robotics and its social impacts (ARSO)*, pp. 125–130. IEEE
69. Złotowski J, Yogeewaran K, Bartneck C (2017) Can we control it? autonomous robots threaten human identity, uniqueness, safety, and resources. *Int J Hum Comput Stud* 100:48–54

**Publisher’s Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.