

Blurring human-machine distinctions

Anthropomorphic appearance in social robots as a threat to human distinctiveness

Francesco Ferrari · Maria Paola Paladino · Jolanda Jetten

Received: date / Accepted: date

Abstract The present research aims at gaining a better insight on the psychological barriers to the introduction of social robots in society at large. Based on social psychological research on intergroup distinctiveness, we suggested that concerns toward this technology are related to how we define and defend our human identity. A *threat to distinctiveness hypothesis* was advanced. We predicted that too much perceived similarity between social robots and humans triggers concerns about the negative impact of this technology on humans, as a group, and their identity more generally because similarity blurs category boundaries, undermining human uniqueness. Focusing on the appearance of robots, in two studies we tested the validity of this hypothesis. In both studies, participants were presented with pictures of three types of robots that differed in their anthropomorphic appearance varying from no resemblance to humans (mechanical robots), to some body shape resemblance (biped humanoids) to a perfect copy of human body (androids). Androids raised the highest concerns for the potential dam-

age to humans, followed by humanoids and then mechanical robots. In Study 1, we further demonstrated that robot anthropomorphic appearance (and not the attribution of mind and human nature) was responsible for the perceived damage that the robot could cause. In Study 2, we gained a clearer insight in the processes underlying this effect by showing that androids were also judged as most threatening to the human-robot distinction and that this perception was responsible for the higher perceived damage to humans. Implications of these findings for social robotics are discussed.

Keywords social acceptance of social robots · threat to human distinctiveness · uncanny valley · robot anthropomorphic appearance · androids

1 Introduction

Technological changes bring innovation but also fears and concerns. From the mechanical innovations in the 19th century to the introduction of computers in the 80s, enthusiasm toward a new technology coincides with suspicion and worries about its possible negative social impact. A similar combination of excitement and concern surrounds the introduction of social robots in today's world. Social robots are designed to interact and communicate with people [1,2] and they vary in terms of capacities and appearance from virtual to humanlike. A recent 2012 Eurobarometer [3] survey into public attitudes toward robots showed that not everyone is unconditionally positively disposed towards this relatively new technology. Whereas the majority (70%) of respondents reported positive attitudes towards robots, many respondents wished to restrict the domains of life where these robots would be used. For example, more than 60% of respondents indicated that it would be inappropriate to utilize these robots to assist in the care and monitoring of the elderly, children and disabled people.

F. Ferrari
Department of Psychology and Cognitive Science
University of Trento
Corso Bettini 31
38068 Rovereto, Italy
Tel.: +39 0464 808625
Fax: +39 0464 808602
E-mail: francesco.ferrari-1@unitn.it

M.P. Paladino
Department of Psychology and Cognitive Science
University of Trento
Corso Bettini 31
38068 Rovereto, Italy
E-mail: mariapaola.paladino@unitn.it

J. Jetten
School of Psychology
The University of Queensland
St Lucia, Qld 4072, Australia
E-mail: j.jetten@psy.uq.edu.au

A more fine-grained analysis suggests that the introduction of social robots leads to questions about human essence and what makes us unique as human beings. For instance, Kamide, Mae, Kawabe, Shigemi, and Arai [4] analyzed the spontaneous comments of 900 Japanese respondents following exposure to a series of videos showing humanoids and androids in everyday life situations. Many comments referred to the fear that these robots would be used for evil and that their presence would threaten human relations, human identity and humanity more generally. Based on these interviews and further quantitative studies, Kamide et al. [4] suggested that, when evaluating social robots, it is important to be mindful of the public's fear that robots endanger humanity and alter human identity (factor named "Repulsion" in the Kamide et al. [4] scale).

Why do people fear that the introduction of social robots will have such a negative impact on humans and their identity? Answering this question would enable us to understand the reasons for resistance to this technological innovation. This would be important because the widespread use of social robots in society at large is only possible when psychological barriers to the introduction of robots in our lives have been removed.

While fear responses can easily be discarded as irrational or caused by people's resistance to change, we argue that social robots pose a specific threat to people. Specifically, social robots, because they are designed to resemble human beings, might threaten the distinctiveness of the human category. According to this *threat to distinctiveness hypothesis*, too much perceived similarity between social robots and humans triggers concerns because similarity blurs the boundaries between humans and machines and this is perceived as damaging humans, as a group, and as altering the human identity. In two studies we put this hypothesis to the test by focusing on robots' anthropomorphic appearance (i.e., the extent to which the robot resembles a human body). In elaborating our predictions, we draw on social robotics' work examining the consequences of robots human-likeness and on social psychological research examining the effect of threat to distinctiveness on intergroup relations. Both lines of research will be reviewed in the next paragraphs.

2 Related work

The *threat to distinctiveness hypothesis* resembles Uncanny Valley theorizing in that it addresses the question why a robot's anthropomorphic appearance may be threatening. In its original version, the uncanny valley theory [5,6] suggests a non-linear relation between a robot's anthropomorphic appearance and its acceptance by humans: human-likeness increases robot familiarity up to a certain point after which further increases in robot human-likeness provoke uneasiness and repulsion in people. According to Ramey [7], the

uncanny feeling is both a cognitive and an affective phenomenon. He suggested that the uncanny feeling evoked by humanlike robots is related to the challenge that these robots pose to the categorical distinction between human and non-humans. For instance, once robots have a human look (e.g., androids), human uniqueness in appearance is undermined. This approach has also been extended to other human characteristics and behavior. For instance, Kaplan [8] suggested that we are afraid of these new machines as they challenge (what we think to be) human uniqueness, forcing us to re-define ourselves and humanness in general. To illustrate the argument, he states that once robots can play chess, the game is no longer thought of as a typically human skill. MacDorman, Vadusevan, and Ho [9] take this reasoning one step further and ask what would happen to our sense of human specialness, if it is possible to create perfect human replicas. For Ramey [7], Kaplan [8] and MacDorman et al. [9], fears and concerns about robots are related to how humans define and defend their identity as human beings. Similarly to the *threat to distinctiveness hypothesis* that we advance in the present research, these authors argue that 'too much similarity of robots to humans' gives rise to fears that this new technology will impact negatively on humans as a group. Note however that despite the fact that there are now a number of studies that have tested uncanny valley theory predictions (e.g., [10,11]), to our knowledge, only MacDorman and Entezari [12] have empirically examined processes related to human-robot distinctiveness. Focusing on the role of individual differences, in a recent correlational study involving a US sample, they found that the extent to which participants conceived of robots and humans as mutually exclusive categories predicted higher feeling of eeriness and lower warmth toward androids. Although the MacDorman and Entezari study underlines the importance of human-robot distinctiveness in the emotional reactions toward robots with a high anthropomorphic appearance, it is worth noting that this study does not provide a direct empirical test of the threat of distinctiveness hypothesis advanced in the present research because robot and human likeness was not manipulated. In addition, and more importantly, these researchers examined the participants' uncanny feelings toward androids. It remains to be seen whether (as examined in the present research) the relationship between robot-human likeness and uncanny feelings map onto concerns about the potential damage to humans and to their identity when robots are introduced into society.

Answering this question is important to understand reasons of societal resistance toward the use and the development of this technology. To do so, we engage with a large body of social psychological work examining the effect of threat to distinctiveness on intergroup relations. Focusing on human groups, studies inspired by the Social Identity Theory [13] have repeatedly shown that people are motivated to see

the social groups they belong to as distinct and different from other groups [13, 14]. By understanding how their own group is different from other groups, group members better understand what makes their group unique (the so called “reflective distinctiveness hypothesis”, [15, 16]). Concerns arise then when there is too much similarity between their own group and another group. Too much intergroup similarity is threatening because it undermines the clarity of intergroup boundaries and challenges that what makes their own group distinctive. One way to cope with this threat is to restore intergroup distinctiveness by differentiating their own group positively from the outgroup (the so called “reactive distinctiveness hypothesis”, see [15, 16]).

We propose that similar processes are at play in relations between humans and robots. As social psychological research on folk conceptions of humanness has shown [17, 18], robots represent a relevant comparison group for humans. Therefore, people tend to spontaneously compare humans with machines to identify core human characteristics. Robots that are able to take on roles typically enacted by humans might thus represent a challenge to the human-machine distinction and therefore their introduction in society is met with greater resistance. Along these lines, in a recent survey investigating hopes and fears toward social robots, Enz, Diruf, Spielhagen, Zoll, and Vargas [19] found that negative attitudes were expressed by respondents who read hypothetical scenarios in which robots were described to have rights equal to humans (i.e., citizenship) or took on roles such as school teacher (e.g., grading the tests of pupils). It remains to be examined whether robots human-like appearance might also represent a challenge to human-machine distinctions.

The *threat to distinctiveness hypothesis* (and the “reactive distinctiveness hypothesis” in particular, see [15, 16]) contributes to a better understanding of why people fear the impact of social robots on human identity and allows us to identify the type of robots that should be most threatening to humans. More specifically, concerns over intergroup distinctiveness would lead us to predict that robots with a highly anthropomorphic appearance — that is those robots that, because of their physical appearance, can be confused with humans — would be the most threatening.

In contrast to industrial mechanical agents, social robots are designed to have a humanlike appearance as this facilitates the use in human-robot interactions of modalities typical of human-human interaction [20, 21, 22]. Typical exemplars of social robots are humanoids such as ASIMO of Honda, HRP-4 of Kawada Industries, NAO of Aldebaran Robotics. However, while humanoids are still quite distinct from humans in terms of physical appearance, as they still have also a mechanical aspect, the same cannot be said of androids whose appearance is designed to be a perfect copy of a human body. Examples of androids are the series of Geminoids (HI-1, HI-2, HI-4, DK, F) created by ATR and Os-

aka University, Philip K Dick and Jules androids of Hanson Robotics, and the FACE robot developed by the FACE Lab of the University of Pisa. Therefore and in line with the distinctiveness threat hypotheses, we expect that for humans, the thought that androids would become part of our everyday life should be perceived as a threat to human identity because this should be perceived as undermining the distinction between humans and mechanical agents.

There is another reason why highly anthropomorphic robots as androids should be perceived as threatening than humanoids. Because of the human-like appearance of anthropomorphic robots and the inability to distinguish them from real humans, such robots could pass themselves off as humans. In other words, they would be able to interact in a human world without being detected and without being recognized for what they really are — and thus they would be impostors. We define impostors in line with a definition put forward by Hornsey and Jetten [23]. An impostor is an individual who publicly claims a group identity (i.e., being vegetarian, being gay, etc.), even if he/she fails to meet all or part of the criteria for group membership (e.g., not eating meat, not having heterosexual relationships). An impostor is thus not a genuine group member, but one who tries to pass as if he/she were, hiding his/her true nature. Jetten, Summerville, Hornsey, and Mewse [24] noted that impostors typically receive very harsh reactions once discovered, especially by members of the group in which they trespass, as they are perceived as damaging the identity of the group they pretend to be part of [25] and because they blur the boundaries between groups [24, 25]. For instance, Warner and colleagues [25] showed that a straight person claiming to be gay was judged by homosexual participants as blurring the boundaries between groups, boundaries that are important for group members as they contribute to self-definition. Even though robots with a highly anthropomorphic appearance may not autonomously decide to pass as a human being, their threat lies in the fact that they have the capability to dilute human identity: it increases the number of those that can appear or act as humans but at the same time it waters down the essence of what means to be human [26].

3 Overview of the research and hypotheses

Given the economic investment in the development of social robots and the likelihood that social robots will increasingly become part of everyday life, it is important to understand the reasons why people fear and resist this development. Several lines of work (reviewed above) suggest that too much similarity between robots and humans threatens the uniqueness of the human category. We predicted that androids (i.e., robots high in anthropomorphic appearance)

in particular should be perceived to threaten intergroup distinctiveness (because they can pass themselves off as humans) are perceived to undermine intergroup boundaries and threaten human identity.

We conducted two studies to test these hypotheses. In both studies (using a between-subjects design in Study 1 and a within-subjects design in Study 2) we presented participants with pictures of three types of robots that differed in their anthropomorphic appearance, varying from no resemblance to humans (mechanical robots), to some body shape resemblance (biped humanoids) to a perfect copy of human appearance (androids). After exposure, we measured the damage that these robots are perceived to cause to humans as a group. We predicted that the perceived damage to humans and their identity would be the highest for androids and the lowest for mechanical robots, with damage perceptions for humanoids in between these two conditions (H1). In addition, in Study 1 we also examined attribution of human qualities and a mind, and predicted, in line with previous findings [27], that mind attribution would be related to the anthropomorphic appearance of the robots, hence to be highest for the android, followed by the humanoid and lowest for mechanical robots. Importantly we expected that robot anthropomorphic appearance, as it elicits a threat to distinctiveness, would be responsible for the perceived potential damage of the robot to human essence and identity (H2).

In Study 2, we aimed to provide a more direct test of the threat to distinctiveness hypothesis asking participants to report to what extent androids, humanoids and mechanical robots were perceived as undermining the human-machine distinction (distinctiveness threat), and their perceived potential damage to humans and human identity. We expected that the perception of undermining human-machine distinctiveness would be highest for the androids and lowest for mechanical robots with treat perceptions for humanoids falling in between these two conditions (H3). Following the threat to distinctiveness account, we predicted that anthropomorphic appearance would elicit the perception that human distinctiveness is undermined (H4a), and this in turn would be responsible for the perception of potential damage to humans and human identity when robots enter into society (H4b).

4 Study 1

4.1 Method

4.1.1 Participants

A total of 182 participants completed all main dependent variables. Participants ($N = 182$, 91 women, 89 men, 2 missing values) were aged between 19 and 63 years ($M_{age} =$

27.70, $SD = 6.36$) and 64% of them reported to have a university degree.

4.1.2 Material: Photos of Robots

In total 18 photos were used to depict 3 mechanical, 3 humanoid, and 3 android robots each with 2 photos (300 pixel width, 400 pixel height). The three mechanical robots were the four legged explorer robot of Toshiba used at the Fukushima Daiichi nuclear implant¹, the Modular Snake robot developed by the Robotics Institute at Carnegie Mellon University called Uncle Sam², and the Nomad Heavy Duty Wheeled Robot of CrustCrawler Robotics³. The three humanoid robots were the HRP-4 developed by AIST and Kawasaki Heavy Industries⁴, the expressive robot Kobian of Waseda University⁵, and the advanced musculoskeletal humanoid robot Kojiro created at the JSK Laboratory at the University of Tokyo⁶. The 3 android robots were the Philip K Dick and Jules robots of Hanson Robotics⁷, the Geminoid DK robots developed by Kokoro for the Aalborg University in Northern Denmark⁸.

The two photos of the mechanical robots depicted the robots from two different points of view. For all androids and humanoid robots, one photo depicted the face of the robot and the other the whole body or the upper part of the body (Jules and Geminoid DK). The most of the pictures were taken from websites of the laboratories that developed the robots, (see footnotes for a complete list). Information on the lab and/or industry that designed the robots was removed from

¹ <http://kmjeepics.blogspot.it/2012/11/toshiba-four-legged-fukushima-robot.html> Retrieved on 25 November 2013;

<http://cdn.phys.org/newman/gfx/news/2012/toshibashows.jpg> Retrieved on 25th November 2013;

² <http://biorobotics.ri.cmu.edu/media/images/fullscreen/snake7.jpg> Retrieved on 25 November 2013;

<http://biorobotics.ri.cmu.edu/media/images/fullscreen/snake5.jpg> Retrieved on 25th November 2013;

³ <http://crustcrawler.com/products/Nomad/index.php> Retrieved on 25th November 2013

⁴ http://www.aist.go.jp/aist_e/latest_research/2010/20101108/20101108.html; AIST: National Institute of Advanced Industrial Science and Technology (of Japan) Retrieved on 25th November 2013

⁵ http://www.takanishi.mech.waseda.ac.jp/top/research/kobian/KOBIAN-R/img/face_movie.jpg; http://www.takanishi.mech.waseda.ac.jp/top/research/kobian/KOBIAN-R/img/2009_neutral.JPG Retrieved on 25th November 2013

⁶ http://h2t-projects.webarchiv.kit.edu/asfour/Workshop-Humanoids2012/kojiro_small.jpg Retrieved on 25 November 2013;

<http://spectrum.ieee.org/image/1534921> Retrieved on 25 November 2013

⁷ <http://www.hansonrobotics.com/robot/jules/> Retrieved on 25th November 2013

⁸ <http://androidegeminoid.blogspot.it/> Retrieved on 25 November 2013;

the photos.

We conducted a pilot study ($N = 24$, 13 women, 10 men, 1 missing value; $M_{\text{age}} = 27.09$, $SD = 2.31$) to check that the androids, humanoids and mechanical robots we had chosen differed in terms of anthropomorphic appearance. Participants were presented with all 18 photos (two for each robot) and were asked “how much does this robot remind you of a human being’s figure?” (responses were recorded on a 7-point scale ranging from 1 “not at all” to 7 “very much”). Subsequently participants were asked to categorize the robots into one of three groups. They were asked to select Group 1 if, in their view, the robot had no or only minimal similarity to humans, Group 2 if the robot was somewhat similar to humans, and Group 3 if the robot was highly similar to humans.

The results of the pilot study are reported in Table 1. We found that mechanical robots were assigned more frequently to the group of robots with minimal or no similarity to human beings (Group 1), humanoid robots to the group of robots that present only some similarity with humans (Group 2), and android robots that are highly similar to humans (Group 3). To further explore these findings, we calculated a mean categorization score for the three groups of robots and submitted this to a one-way repeated measures ANOVA (robot: mechanical vs. humanoid vs. android). Least Significant Difference (LSD) was used as post-hoc comparison test. Mauchly’s test revealed that the assumption of sphericity was violated, $\chi^2(2) = 17.711$, $p < .001$, and we therefore corrected the degrees of freedom (DoF) using Greenhouse-Geisser estimates of sphericity ($\epsilon = .64$). A significant main effect was found, $F(1.29, 29.62) = 299.95$, $p < .001$, showing that androids ($M = 6.51$, $SD = .57$) were perceived as most similar to humans, followed by humanoids ($M = 3.76$, $SD = 1.39$), and then by mechanical robots ($M = 1.10$, $SD = .18$), all $ps < .001$.

4.1.3 Material: Human nature traits

Forty traits were used to measure *Attribution of Human Nature traits* toward the robots. These traits were chosen on the basis of a pilot study in which participants ($N = 48$, 32 females, 16 males; $M_{\text{age}} = 24.83$, $SD = 3.8$) were asked to rate 71 traits on the two dimensions of Humanness identified by Haslam [17]. Specifically, questions were included to assess human nature (“Is this feature typical of the Human Nature, as it makes us human and therefore different from machines”) as well as human uniqueness (“To what extent each of the following characteristics is uniquely human, and therefore is not present in other animal species?”). We also assessed the valence of the trait (“Indicate for each trait to what extent it is, in your opinion, positive or negative”), and the appropriateness of the trait to describe a robot (“Would you use this feature to describe a robot, its functions and be-

havior?”). From the 71 traits, we selected 20 traits high in human nature and 20 traits low in human nature that were equivalent in terms of valence, $t(47) = -.425$, $p > .05$, and that did not differ in terms of uniquely humanness, $t(25) = -.337$, $p > .05$. In addition, all selected items were judged to be appropriate to describe robots (see table 2).

4.1.4 Procedure

Participants were contacted via-email and Facebook and invited to participate to an on-line study assessing people’s opinions of robots. Participants were informed that data collection would be anonymous, that their responses would remain confidential and that they had the right to withdraw from the study at any stage without penalty. Once consent was obtained, participants were directed to a questionnaire showing pictures of one robot. Robot anthropomorphic appearance was manipulated between-subjects (androids vs. humanoids vs. mechanical robots). After viewing the pictures, participants completed a questionnaire including, among others, measures that are of interest to test the threat to distinctiveness hypothesis.

4.1.5 Dependent Variables

We relied on the work of Kamide et al. [4], and used items of the Psychological Scale for General Impressions of Humanoids (PSGIH), when relevant, to measure the constructs under investigation. To our knowledge, the work of Kamide et al. [4] represents the only attempt in the field of social robotics to examine in a bottom-up way (i.e., starting with interviews followed by questionnaire, etc.) the evaluation of social robots on different dimensions. This work has resulted in a set of items that can be used to quantify these evaluations (see [28]).

4.1.6 Anthropomorphic and Robotic Appearance

An index of robot *anthropomorphic appearance* was created by averaging responses to the following three items: “I could easily mistake the robot for a real person”, “The robot looks like a human”, “I think the robot looks too much like a human” ($\alpha = .88$). We created another index of Robotic Appearance averaging the responses to the items: “I do not get the impression that it is a robot at all when I look at it” (reverse scaled), “The robot looks like a robot”, and “The robot is like a robot in every way” ($\alpha = .85$). In the original PSGIH scale [4] (see also [28]) these items loaded on the same factor (labeled “Humanness”). Given that our hypotheses concern robot anthropomorphic appearance and not the

Table 1 Results of pilot study on anthropomorphic appearance of robots used in study 1. Means of how much each robot reminds the figure of a human being (second column), and the percentage with which it is associated to each group (third, fourth and fifth columns) are reported.

Type of robot	$M(SD)$	Group 1 Minimal or no similarity to humans	Group 2 Some Similarity to Humans	Group 3 Very similar to humans
PkD	6.79 (.51)	-	-	100%
Gemindoid DK	6.75 (.61)	-	-	100%
Jules	6.00 (.98)	-	8.3%	91.7%
Kojiro	3.92 (1.56)	4.2%	95.8%	-
HRP-4	3.74 (1.66)	8.3%	91.7%	-
Kobian	3.67 (1.49)	4.2%	95.8%	-
Toshiba fourlegged	1.21(.42)	95.8%	4.2%	-
Nomad Heavy Robot	1.04 (.20)	100%	-	-
Uncle Sam	1.04 (.20)	100%	-	-

Table 2 Traits high and low in human nature

High Human Nature	Low Human Nature
Aggressive	Accurate in reasoning
Ambitious	Active
Childish	Analytic
Comfortable	Cold
Conscientious	Competent
Determined	Conservative
Easily distractible	Disinterested (no ulterior motives)
Friendly	Do the things automatically
Frivolous (fatuuous)	Hard-hearted
Impatient	Ignorant
Impulsive	Unsophisticated (simple-minded)
Irresponsible (does not want to take responsibility)	Passive
Judicious	Rational
Not self-confident	Refined mentality
Pleasant from an interpersonal perspective	Reliable (of which you can be trusted)
Pleasant	Repetitive
Rude	Shallow
Sensible	Skillful
Sympathetic	Strict
Wary	Unable to collaborate

robotic appearance (see also the result session), we kept these set of items separate⁹.

4.1.7 Damage to Humans and to Human Identity

Four items were used to assess perceived damage of robot on humans and their identity: “The robot seems to lessen the value of human existence”, “I get the feeling that the robot could damage relations between people”, “The robot could easily be used for evil (to fool, to harm, etc.)” and “I think the robot will soon control humans”. Responses to these items were averaged to create an index of *damage to humans and to human identity* ($\alpha = .78$). In the original PS-GIH scale, these items concerning the potential social damage of robots loaded in the so-called “Repulsion - anxiety toward the existence of robots” factor.

⁹ This factor also included an item assessing human qualities attributed to robots. This item will not be considered further as it is not relevant to assess support for the current hypotheses.

Responses for Anthropomorphic appearance, Robotic appearance and Damage to Humans were recorded on a 7-point Likert Scale with values ranging from 1 = “strongly disagree”, to 4 = “neither agree or disagree”, to 7 = “strongly agree”.

4.1.8 Mind and Human Nature traits attribution

Participants were asked to what extent the robots seemed like to have the following *mind experience* and *mind agency* capacities: fear, pain, pleasure, joy (for mind experience) and planning, emotion recognition, self-control, morality (for mind agency). An example item is: “it seems like this robot can feel pain”. These capacities were chosen on the basis of a factor analysis by Gray, Gray and Wegner [29] confirming these items capture the two types of minds. An index of mind experience attribution (average of the items’ responses; $\alpha = .95$), and another for mind agency (average of the items’ responses; $\alpha = .71$) were created. Responses

were recorded on a 7 point Likert Scale ranging from 1 “not at all” to 7 “completely”.

Participants were asked to what extent each of the twenty traits high and the twenty traits low in human nature were descriptive of the robot (To what extent does this feature describe the robot in the picture?). The order of presentation of the traits was randomized for each participant. Participants recorded their answers on a 7 point Likert Scale (1 not at all to 7 very much). The responses to the 20 *high human nature* ($\alpha = .89$) and the 20 *low human nature traits* ($\alpha = .85$) were averaged to create an index of high human nature and an index of low human nature robot attribution.

At the end of the questionnaires we asked participants to indicate their age, sex, education, and the device they used to respond to the questionnaire. Finally, participants were presented with a debrief and an email address in case they would like further information.

4.2 Results

Preliminary analysis including sex of the participants showed that this variable influenced the results for the following dependent variables: ratings of anthropomorphic appearance, damage to humans and to human identity, mind agency and high human nature traits attribution. These variables were analyzed in a Robot (mechanical vs. humanoid vs. android) \times participants' sex between subjects ANOVA. For all other analyses, data were submitted to a one-way between subjects ANOVAs (Robots: mechanical vs. humanoid vs. android), and Least Significant Difference (LSD) were used as the post-hoc comparison test following up significant effects. The results for all dependent variables are presented in table 3^{10,11}.

4.2.1 Anthropomorphic and Robotic Appearance

An ANOVA revealed an effect of Type of Robot on *Anthropomorphic Appearance*, $F(2, 174) = 201.87, p < .001$, indicating that our manipulation was successful. Androids were

¹⁰ Part of these data were also used in Ferrari and Paladino (2014) — a study that focused on validating the scale developed by Kamide and colleagues in an Italian sample.

¹¹ In Study 1, participants were also asked to record their highest level of education to date ($N = 3$ ‘secondary school’, $N = 60$ ‘high school’, $N = 32$ ‘bachelor degree’, $N = 68$ ‘master degree’, $N = 16$ ‘Phd or superior degree’, and 3 missing). Exploratory analyses were conducted exploring the role of educational level on the two main dependent variables of Study 1: robot anthropomorphic appearance and damage to humans. Specifically, in the ANOVAs, participants level of education was included as a covariate or as a factor (recoded whereby 0 = high school degree or lower, $N = 63$; 1 = university degree or higher, $N = 116$). No significant effects were obtained for level of education and results for anthropomorphic appearance (all $ps > .16$) and for damage to humans and their identity (all $ps > .55$) were unaffected by inclusion of education in the analyses.

judged as most similar to human beings ($M = 4.91, SD = 1.34$), followed by humanoids ($M = 2.15, SD = 1.1$) and then by mechanical robots ($M = 1.22, SD = .62$), all comparisons, $ps < .001$. Interestingly, the type of robot \times participant sex interaction was significant, $F(2, 174) = 3.09, p = .05$, showing that male and female participants differed in how they judged androids and humanoids appearance. Androids tended to be rated as more human-like by female ($M = 5.16, SD = 1.22$) than male participants ($M = 4.68, SD = 1.43$), $F(1, 174) = 3.11, p = .08$, whereas humanoids were judged as slightly more human-like in appearance by male ($M = 2.38, SD = 1.13$) than female participants ($M = 1.92, SD = 1.01$), $F(1, 174) = 2.83, p = .10$. It is worth noting that both effects were only marginally significant, and more importantly, that the interaction did not alter the success of our manipulation. Indeed, when examining effects separately for male and female participants, we found the Type of Robot main effect both for female, $F(2, 88) = 126.16, p < .001$, and male participants, $F(2, 86) = 81.58, p < .001$. Both male and female participants judged androids as most human-like, followed by humanoids, and then mechanical robots (all mean comparisons $ps < .02$). The three types of robots were also differently judged in terms of *Robotic Appearance*, $F(2, 179) = 86.63, p < .001$. Interestingly, in terms of appearance, Humanoids ($M = 6.23, SD = .91$) were judged as the most typical robots, followed by the mechanicals ($M = 5.07, SD = 1.73$) and finally by androids ($M = 2.99, SD = 1.33$), all $ps < .001$.

Taken together, these results show that androids were perceived as the robots that resembled humans most and robots least. Interestingly, humanoids were judged most robotic in appearance suggesting that, in participants mind, this kind of robots maps best onto the mental “robot” schema.

4.2.2 Damage to Humans and to Human Identity

Consistent with H1, perceived damage to humans and to their identity differed by condition, as indicated by the type of robot main effect, $F(2, 174) = 9.00, p < .001$. Specifically, androids were judged as potentially more damaging ($M = 3.23, SD = 1.51$) than humanoids ($M = 2.62, SD = 1.32$) and more damaging than mechanical robots ($M = 2.19, SD = 1.28$), all $ps < .01$. Humanoids were perceived as marginally significantly more damaging than mechanical robots, $p = .08$. The main effect for participant sex was also significant, $F(1, 174) = 5.68, p < .02$, highlighting that females were more concerned about robots ($M = 2.91, SD = 1.54$) than males ($M = 2.44, SD = 1.27$).

4.2.3 Mind Attribution

We also found that mind attribution was influenced by type of robot, $F(2, 177) = 10.45, p < .001$. Mind experience was

Table 3 Means of the androids, humanoids and mechanical robots for the different dependent variables of study 1

Type of Robot	Anthropomorphic Appearance	Robotic Appearance	Damage to Humans and Their Identity	Mind ence	Experi- ence	Mind Agency	High Human Nature Traits
Mechanical	1.22 ^a	5.07 ^a	2.19 ^a	1.35 ^a		2.50 ^a	2.15 ^a
Humanoid	2.15 ^b	6.23 ^b	2.62 ^a	1.80 ^b		2.89 ^{ab}	2.57 ^b
Android	4.91 ^c	2.99 ^c	3.23 ^b	2.39 ^c		3.17 ^b	2.86 ^b

Values in each column with different subscript are significantly different from each other at $p < .05$.

attributed most to androids ($M = 2.39, SD = 1.58$), followed by humanoids ($M = 1.80, SD = 1.22$), and by mechanical robots ($M = 1.35, SD = .84$), all comparisons were significant, $ps = .05$.

For mind agency attribution, a main effect of type of robot, $F(2, 174) = 4.47, p < .02$ emerged. Mechanical robots ($M = 2.50, SD = 1.16$) were attributed less mind agency than androids ($M = 3.17, SD = 1.37$), $p < .005$, and (albeit only marginally significantly so) less mind agency than humanoids ($M = 2.89, SD = 1.24, p = .09$). Androids and humanoids were not significantly different from each other, $p > .22$. However, this main effect was qualified by an interaction between robots and participants sex, $F(2, 174) = 3.43, p < .04$. Separate one-way ANOVAs for male and female participants showed that this tendency was only significant for male participants, $F(2, 86) = 7.23, p < .002$. Mind agency characterized androids ($M = 3.35, SD = 1.45$) and humanoids ($M = 3.16, SD = 1.21$, not significantly different from each other $p > .54$), more so than mechanical robots ($M = 2.21, SD = 1$), all comparison with mechanical robots, $ps < .005$. In contrast, for female participants, there were no differences between conditions, $F(2, 88) = .64, p > .52$.

4.2.4 Human nature traits attribution

Analysis of high human nature traits attribution revealed a main effect of type of robot, $F(2, 174) = 9.09, p < .001$. Androids ($M = 2.86, SD = .90$) were judged to possess these traits to a greater extent than mechanical robots ($M = 2.15, SD = .87$), $p < .01$, and only marginally significant more so than humanoids ($M = 2.57, SD = .97$), $p = .08$. Humanoids were judged to possess high human nature traits to a greater extent than mechanical robots, $p < .02$. There was also a marginal significant effect of participant sex, $F(1, 174) = 3.59, p = .06$, showing the tendency for females ($M = 2.40, SD = .92$) to attribute fewer high human nature traits to robots compared to males ($M = 2.66, SD = .98$).

An ANOVA revealed no main effect of type of robots on low human nature traits, $F(2, 177) = 2.20, p > .11$.

4.2.5 Testing the role of anthropomorphic appearance on perceived damage to humans and their identity: Mediation analysis

The results suggested a linear pattern between the increase of robots' anthropomorphic appearance and the perceived damage to humans and their identity. To further explore this finding, we conducted additional analyses to verify whether ratings of anthropomorphic appearance mediated the effect of robots on perceived damage to humans and their identity ($N = 182$). All the analyses were conducted with INDIRECT, a macro for SPSS provided by Preacher and Hayes [30].

We first regressed the potential mediator (anthropomorphic appearance), and then the dependent variable (damage to humans and their identity), on our independent variable: type of robot (coded as continuous variable, Mechanic = 0, Humanoid = 1, and Android = 2; see [31,32,33] for similar approach). In line with the previous analysis, these regressions showed a significant effect both on anthropomorphic appearance ($b = 1.84, SE = .11, t(180) = 17.56, p < .001$), and damage to humans and their identity ($b = .51, SE = .12, t(180) = 4.12, p = .001$). Subsequently, we regressed damage to humans and their identity simultaneously on anthropomorphic appearance and type of robot, and found that anthropomorphic appearance was positively associated with the dependent variable ($b = .39, SE = .08, t(180) = 4.63, p < .001$).

We tested the overall significance of mediation using the bootstrap method recommended by Fritz and MacKinnon [34]. For this analysis, the 95% confidence interval of the indirect effect was obtained with 5000 bootstrap resamples. We constructed bias-corrected confidence intervals around the product coefficient of the indirect (mediated) effect using the SPSS macro Preacher and Hayes [35] created. The product coefficient is based on the size of the relationship between the independent variable and the mediator and the relationship between the mediator and the dependent variable. The indirect effect was .71, with a confidence interval ranging from .32 to 1.2. Because the confidence interval does not include zero, the indirect effect was significant. Finally, the analyses indicated that the direct effect of robots on perceived damage to humans and their identity did not reach significance, when controlling for ratings of anthropomor-

phic appearance ($b = .20, SE = .19, t(180) = 1.04, p > .3$) — a pattern of results suggestive of full mediation.

Exploratory, we also investigated whether the attribution of mind experience or the attribution of traits high in human nature mediated the effect of type of robots on perceived damage to humans and their identity. Consistent with the ANOVA, mind experience and high human nature traits, were significantly affected by type of robot, all $ps > .001$. However, when simultaneous regressing perceived damage to humans and their identity on type of robot and mind experience, this latter variable was not significant ($p > .19$) suggesting that mind experience was not responsible for the effect of damage on type of robot. Likewise, there was no evidence that attribution of traits high in human nature mediated this relationship ($p > .19$).

4.3 Discussion

To sum up, consistent with H1 we found that androids - whose appearance is modeled on that of a human body - raised the highest concerns for the potential damage to humans and human identity, followed by humanoids and then mechanical robots. Importantly, and consistent with H2, the mediation analysis demonstrated that robot anthropomorphic appearance, and no other aspects on which the three types of robots differed (i.e., the attribution of mind and human nature traits), was responsible for the perceived damage that the robot could cause to humans and their identity. All in all, these findings are consistent with the idea that worries and concerns about the impact on human identity of highly human-like social robots are related to the fact that these robots look so similar to humans that they can be mistaken to be one of us.

5 Study 2

The aim of Study 2 was twofold. We aimed to replicate Study 1 findings and also sought to test the threat of distinctiveness hypothesis more directly. To do this, in addition to perceived anthropomorphic appearance and perceived damage to humans and human identity, participants were also asked to rate to what extent they perceived that androids, humanoids and mechanical robots were undermining the categories distinction between machines and humans. Following our threat to distinctiveness hypothesis, we expected a similar pattern of results on the perception of damage to humans and their identity (H1) as on a blurring of human-machine distinction measure (H3): androids should be perceived as most likely to blur boundaries, followed by humanoids and then mechanical robots. We also examined if anthropomorphic appearance elicits the threat to human distinctiveness, operationalized as the perception that the human-machine

distinction is undermined (H4a), and whether distinctiveness threat is responsible for the perceived potential damage of the robot to human essence and identity (H4b). These hypotheses were tested in a within-subjects design.

5.1 Method

5.1.1 Participants

Fifty-one participants (49 females and 2 males) aged between 19 and 23 years ($M_{age} = 20.2, SD = .67$) completed the questionnaire. Participants were all students of the Department of Psychology and Cognitive Science of University of Trento, and they received credits for their participation.

5.1.2 Material

Two pictures each (97 pixel for width x 130 pixel for height) for 4 Mechanic, 4 Humanoid and 4 Android robots (a total of 24 images) were used. The pictures were the same as used in Study 1, with a few exceptions. In the mechanical robot group, the photos of snake robot Uncle Sam were substituted with those of WowWee's Rovio¹². In addition, we added the pictures of the tracked robot "TP-600-270"¹³ developed by SuperDroid Robots. For humanoids, instead of HRP-4, we used photos of Wabian-2 of Waseda University¹⁴, and those of Tichno R of V-Stone¹⁵. Finally for the android group, in addition to the photos used in Study 1, we added two images of FACE android developed by FACE Lab of University of Pisa [36,37]. Similar to Study 1, for mechanical robots, each photo depicted the robot from two different points of view, whereas for humanoid and android robots, one photo depicted the face of the robot and the other the whole body or the upper part of the body (Jules, Geminoid DK, and FACE). Most pictures were selected from websites of the laboratories that developed the robots, (see footnotes for a complete list) with the exception of the photos of the FACE android. These photos were made available by the FACE Lab. As in

¹² <http://blog.tmcnet.com/blog/tom-keating/gadgets/rovio-wi-fi-voip-robotic-webcam.asp>. Retrieved on 25 November 2013;

¹³ <http://www.superdroidrobots.com/shop/item.aspx/new-prebuilt-hd2-s-robot-with-5-axis-arm-and-cofdm-ocu-sold/1279/>. Retrieved on 25 November 2013;
http://www.superdroidrobots.com/product_info/UGV\%20System\%20Design.pdf. Retrieved on 25 November 2013

¹⁴ http://www.takanishi.mech.waseda.ac.jp/top/research/wabian/img/wabi_front2008.jpg Retrieved on 25 November 2013

¹⁵ <http://www.sansokan.jp/robot/showroom/11.html> Retrieved on 25 November 2013;
<http://www.zimbio.com/pictures/p1UE1otXSWW/Robot+Venture+Companies+Hold+Joint+Press+Conference/KF3TfpVxLcD/Vstone+Tichno> Retrieved on 25 November 2013

Study 1, information on the labs and/or industries that developed the robots were removed from the photos.

5.1.3 Procedure

Participants, in groups of maximum 10 people, completed the online questionnaire in one of the university lab. After reading and signing the informed consent, they were invited by the experimenter to start the study. The study was presented as an investigation of opinions toward different kinds of robots. At the beginning participants were asked to indicate their age, sex, education and occupation and then they were asked to complete the Humanity Esteem Scale [38]¹⁶. Then, pictures of all robots were presented on a single page, and participants were informed that all robots were real robots, developed by different laboratories in the world. In the following pages, participants were asked to complete, among others, the scales on physical anthropomorphism, threat to human machines boundaries and damage to humans and their identity (and other items that will not be considered here) for androids, then for humanoids and finally for mechanical robots (the order of robots presentations and questions was randomized across participants). All items were presented next to the photos of the robots so that the pictures were always visible.

5.1.4 Dependent Variables

If not further specified responses were recorded on a 7-point Likert scale, (1 = “strongly disagree”, 2 = “moderately disagree”, 3 = “slightly disagree”, 4 = “neither agree or disagree”, 5 = “slightly agree”, 6 = “moderately agree”, 7 = “strongly agree”).

5.1.5 Anthropomorphic appearance

The same items as used in Study 1 were included. As before, an index (average of the responses) for androids ($\alpha = .74$), humanoid ($\alpha = .60$) and mechanical robots was calculated for each participant. The Cronbachs alpha was not calculated for the mechanical robots because there was limited variability in the responses.

5.1.6 Undermining human-machine distinctiveness

The following three items were used to assess this construct: “This type of robot gives me the impression that the differences between machines and humans have become increasingly flimsy”, “Looking at this kind of robot I wonder/ask myself what are the differences between robots and

humans”, and “This type of robot blurs the boundaries between human beings and machines”¹⁷. These were adapted from the study of Warner et al. [25]. A mean score was calculated for this undermining human-machine distinctiveness measure – for androids ($\alpha = .83$) for humanoids ($\alpha = .62$) and for mechanical robots ($\alpha = .36$).

5.1.7 Damage to humans and their identity

We used the same four items used in study 1. The mean damage to humans and their identity score was calculated for each participant separately for mechanical robots ($\alpha = .59$), humanoids ($\alpha = .72$), and androids ($\alpha = .70$).

5.2 Results

If not further specified, the data were analyzed in one-way repeated measures ANOVAs (Robots: mechanical vs. humanoid vs. android) and the Least Significant Difference (LSD) was used as post-hoc comparison test. The results for the dependent variables are described below and shown in table 4.

5.2.1 Anthropomorphic appearance

Mauchly’s test revealed that the assumption of sphericity was marginally violated, $\chi^2(2) = 5.69$, $p = .058$, therefore DoF were corrected using Huynh-Feldt estimates of sphericity ($\epsilon = .93$). As in Study 1, the main effect was significant, $F(1.87, 93.28) = 584.62$, $p < .001$, showing that androids were rated as physically most similar to human beings ($M = 5.97$, $SD = 1$), followed by humanoids ($M = 2.03$, $SD = 1$) and then by mechanical robots ($M = 1.07$, $SD = .29$), all $ps < .001$.

5.2.2 Undermining human-machine distinctiveness

Mauchly’s test revealed that sphericity was partially violated, $\chi^2(2) = 5.9$, $p = .052$, therefore we corrected DoF using Huynh-Feldt estimates of sphericity ($\epsilon = .93$). There was a significant effect, $F(1.86, 92.95) = 90.4$, $p < .001$, showing that androids were perceived as the robots that blurred the distinctiveness between human and machines to the greatest extent ($M = 4.47$, $SD = 1.61$), followed by humanoids ($M = 2.72$, $SD = 1.26$) and then by mechanical robots ($M = 1.73$, $SD = .82$), all $ps < .001$.

¹⁶ Exploratory analysis indicated that Humanity Esteem did not moderate any of the findings. For the sake of brevity, these results are therefore not presented

¹⁷ Initially there was a fourth item (“This type of robot highlights that there are clear differences between humans and machines”) that we excluded it to increase the reliability of undermining to human-machine distinctiveness scale.

Table 4 Means and Standard Deviation (in parenthesis) for the dependent variables and the different kinds of robots of study 2.

Type of Robot	Anthropomorphic Appearance	Undermining Human-Machine Distinctiveness	Damage to Humans and Their Identity
	<i>M(SD)</i>	<i>M(SD)</i>	<i>M(SD)</i>
Mechanical	1.07 (.29)	1.73 (.82)	2.78 (1.09)
Humanoid	2.03 (1)	2.72 (1.26)	3.08 (1.27)
Android	5.97 (1)	4.47 (1.61)	4.16 (1.28)

Values in each column are significantly different from each other at $p < .05$.

5.2.3 Damage to humans and their identity

Mauchly's test was not significant, $\chi^2(2) = .944, p > .24$, and sphericity not violated. Type of robot revealed a significant effect, $F(2, 100) = 65.72, p < .001$. As in Study 1, Androids ($M = 4.16, SD = 1.28$) were perceived as the robots that were most likely to negatively affect humans, followed by humanoids ($M = 3.08, SD = 1.27$) and by mechanical robots ($M = 2.78, SD = 1.09$), all $ps < .015$.

5.2.4 Anthropomorphic appearance, undermining human-machine distinctiveness, damage to humans and their identity: Mediation analysis

The results suggest a linear pattern for the increase of robots' anthropomorphic appearance, undermining human-machine distinctiveness and perceived damage to humans and their identity. Further analysis were conducted to test the role of anthropomorphic appearance of the type of robots on undermining human-machine distinctiveness (H4a) and then the possible mediation of undermining human-machine distinctiveness on the relation between type of robots and damage to humans and their identity (H4b). To this end, we conducted two separate analyses following the approach of *causal steps* [39, 40, 41]. Through this approach we observed if the effect of kind of robot (factor) on the dependent variable (first undermining human-machine distinctiveness and then damage to humans and their identity), was reduced when the mediator (first anthropomorphic appearance and then undermining human-machine distinctiveness) was included into the analysis/equation. A significant effect of the mediator is suggestive of mediation. We analyzed the data using the Linear Mixed Model (LMMs) procedure in SPSS. If not further specified we selected a first order autoregressive (AR1) covariance structure in our repeated measures analyses, which assumes that residual errors within each subject are correlated but independent across subjects. Intercepts and participants were entered in the model as random effect.

We tested first the mediation of anthropomorphic appearance on undermining human-machine distinctiveness. When entered as a repeated measure fixed effect, in line with the

previous analysis (ANOVAs), we found that type of robot significantly affected undermining human-machine distinctiveness (dependent variable), $F(2, 73.78) = 79.004, p < .001$, and anthropomorphic appearance (proposed mediator), $F(2, 68.38) = 530.893, p < .001$. In a further LMMs analysis, anthropomorphic appearance (covariate) was entered as repeated measure fixed effect and we found that it significantly affected undermining human-machine distinctiveness, $F(1, 68.34) = 244.604, p < .001$. Finally we entered simultaneously type of robot (independent variable) and anthropomorphic appearance (covariate) as fixed effects. We found significant effects for both anthropomorphic appearance, $F(1, 146.13) = 43.692, p < .001$, and type of robot, $F(2, 89.98) = 4.581, p < .05$. However, it is worth noting that the influence of type of robot was strongly reduced when we included anthropomorphic appearance in the equation confirming its role as mediator of the effect of type of robot on undermining human-machine distinctiveness. This pattern of data suggests that robots human-likeness directly increases the perception of robot as a source of danger to humans and their identity: the more the robot's appearance resembles that of a real person, the more the boundaries between humans and machines are perceived to be blurred.

We then tested whether undermining human-machine distinctiveness mediates the effect on damage to humans and their identity. In line with the previous analysis (ANOVAs), we found that type of robot entered as a fixed effect significantly affected damage to humans and their identity, $F(2, 63.89) = 55.465, p < .001$. Next, we entered undermining human-machine distinctiveness (covariate) as a fixed factor, and we found a significant effect on damage to humans and their identity, $F(1, 88.97) = 73.13, p < .001$. A further LMMs analysis was conducted entering simultaneously type of robot (independent variable) and undermining human-machine distinctiveness (covariate) as fixed factors and damage to humans and their identity as the dependent variable. The results showed that both the effect for undermining human-machine distinctiveness, $F(1, 124.693) = 6.221, p < .015$, and type of robot, $F(2, 74.028) = 14.769, p < .001$, were significant. However, when we included undermining human-machine distinctiveness in the equation, the influence of type of robot was reduced. Even though the effect of type of robot was still significant, the results are suggestive of mediation by undermining human-machine distinctiveness: highly anthropomorphic robots, such as androids, are perceived as damaging humans and their identity because they blur the boundaries between machines and human beings, undermining the sense of *being human* (see Fig. 1).

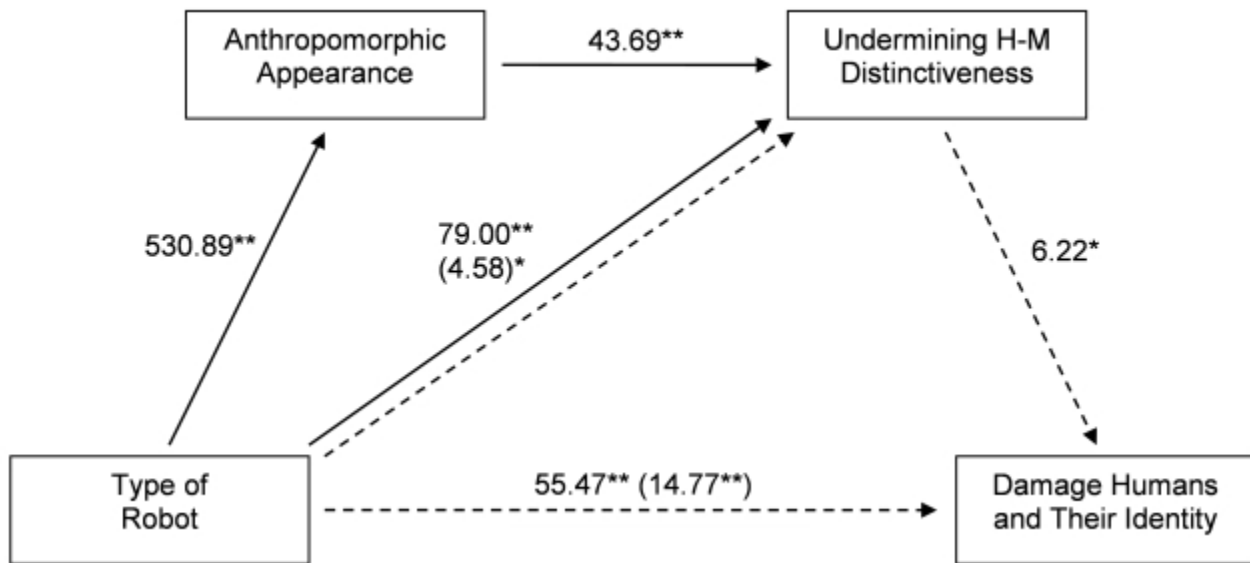


Figure 1 Representation of mediation effects between Type of Robot factor, Anthropomorphic Appearance, Threat to Distinctiveness and Damage to Humans. The continuous arrows indicate the first mediation analysis between Type of Robots, Anthropomorphic Appearance (Mediator 1), and Undermine Human-Machine Distinctiveness. The dotted arrows describe the second mediation analysis between Type of Robots, Undermine Human-Machine Distinctiveness (Mediator 2), and Damage to Humans and Their Identity. We reported the F values of LMMs analysis for each relation and indicated in parentheses the F values of Type of Robot factor controlling for the mediators. *= $p < .05$; **= $p < .001$;

5.3 Discussion

Consistent with H1 and the findings of Study 1, in Study 2, a clear linear effect emerged on all measures, showing that androids were rated as most anthropomorphic, most of a threat to the distinction between humans and machines and most damaging to humans as a group, and to their identity (followed by humanoids and mechanical robots). Note that even though androids also elicited highest concerns for the potential damage to humans and their identity in Study 1 that linear relationship was not observed on all measures. One reason for this difference may be that in Study 2 a within-subjects design was used whereby each participant saw and judged every type of robots. This methodological design has the advantage over a between-subjects design in that it better controls for individual differences, and maximizes comparisons between robots. Both aspects could have contributed to the finding that the differences among these three types of robots are more clear-cut in Study 2 compared to Study 1. In addition, in this study we gained a clearer insight in the underlying processes. The mediational analyses showed that the ratings of robot anthropomorphic appearance was responsible for the differences in the perception of undermined human-machine distinctiveness (confirming H4a). In turn, judgments of undermined human-machine distinctiveness accounted for the differences in the perceived robots damage to humans and their identity (confirming H4b). All in all, these findings are consistent with a *threat to distinctiveness hypothesis*: participants fear highly anthropomorphic robots

(i.e., robots that look too similar to humans), as they blur the distinction between humans and mechanical agents.

6 General Discussion

In the present research we aimed to gain a better insight in the question why people fear the introduction of social robots in their daily life. Based on works of Ramey [7], Kaplan [8], MacDorman et al. [9,12], and intergroup distinctiveness research [15,16], we suggested that concerns toward the negative impact of the entering of this technology in our life is related to how we define and defend our human identity. Specifically, we advanced the *threat to distinctiveness hypothesis* suggesting that too much similarity between robots and humans gives rise to concerns that the distinction between humans and mechanical agents is blurred, thereby threatening intergroup distinctiveness. In two studies we tested and found support for this hypothesis observing participants reactions to three types of robots that varied from low (i.e., mechanical robots) to medium (i.e., humanoids) to high anthropomorphic appearance (i.e., androids). The findings of the present research have some important implications for social robotics research and specifically for how a robot's appearance affects reactions to robots. The findings suggest that one way to improve robots' acceptance is to increase robot familiarity. With this goal in mind, roboticians have developed humanlike robots as they are supposed to elicit responses and behaviors typically shown towards human partners [42,43]. Our research suggests that this goal

should however not conflict with “the need for distinctiveness” that typically characterizes intergroup comparisons. Indeed, and as we show here, such concerns extend to humans-robots relations. Robots are more likely to be accepted when differences and distinctiveness from human beings is somehow preserved. In this regard, it should be noted that according to the threat to distinctiveness hypothesis the factor that triggers concerns is not robot-human similarity *per se*, but “too much” similarity which blurs the boundaries between humans and mechanical agents. In the present research, only highly anthropomorphic android robots reached this point. Differently from humanoids and industrial robots, androids (who are built to be perfect copies of human bodies with no visible mechanical elements) were on average judged as “looking too much like a human” and “as easily mistaken for one of us” (see the scores of the anthropomorphic appearance ratings in both studies). At the same time, the introduction of these robots in society was also judged on average as having a negative impact on humans as a group. In this regard, the present research provides empirical support to one of the guidelines proposed by the project “RoboLaw”. Funded by the EU, the goal of this research project was to promote a technically feasible, and ethically and legally sound basis for future robotics developments (<http://www.robolaw.eu>). According to the researchers, one way to reach this goal is to avoid that a robot, including its appearance, could deceive people.

The present findings also have interesting implications for the uncanny valley theory and more generally for theoretical work on the effects of robot-human likeness. According to Ramey [7], emotional reactions toward androids are related to the fact that they challenge the categorical distinction between humans and machines. Consistent with this, MacDorman and Entenzari [12] showed that the extent to which humans and robots were considered to be highly distinctive categories (measured as an individual difference) predicted uncanny feelings towards androids. In the present research we extend this finding by showing that distinctiveness is also key to understanding resistance to the introduction of these robots in society. Indeed, we found that androids (compared to humanoids and mechanical) were most likely to be seen to undermine the distinctiveness between humans and robots.

The findings of our research also provide empirical support for Ramey’s theorizing [7] that androids represent a problem for the way we, as humans, define and defend our identity when presented with highly humanlike robots. Consistent with this, we showed that concerns about androids are similar to those typically registered when responding to impostors: the fear that these individuals could alter the group’s identity [23, 24, 25]. Finally, drawing a link between responses toward social robots and responses to other type of threats, our research underlines the importance to engage with social psychological theorizing on intergroup relations

when designing and evaluating the impact of social robots (for other examples of studies in social robotics relying on intergroup relations theorizing, see also [44, 45]).

Our findings also help to understand societal resistance toward the introduction of social robots in society, providing a better insight in the question *why* people do or do not fear the use of social robots. Previous studies have shown that social beliefs concerning a technology play an important role. These beliefs can have a direct and an indirect influence (through social influence on what important others think) on its acceptance. For instance, willingness to use assistive social agents technology (e.g., RoboCare robot) among older adults depends also on the perceived consequences of the use of that technology. If these are positive (i.e., the robot would make life more interesting) and are shared by important others, it has been found that the intention of older adults to use the robot significantly increases [46]. In this line of reasoning, our research suggest that robots that do not challenge the human-machine distinctiveness are more likely to be sought out, used and recommended to others. This turns us to the question of how to design robots that evoke familiar responses and, at the same time, do not challenge human’s need for distinctiveness. This will be discussed next.

7 Limitations and Future Research

As every research, the present one also has some limitations. The most obvious is that we used photos and not videos or direct interactions with robots. Note however that this methodology is common and also used in other studies investigating the role of robot appearance (e.g., [47]) as it allows for optimal control (e.g., no interference with a robot’s movement ability). That said, future studies on the societal resistance to the development of robots should also consider more complex and richer materials and contexts. Compared with just viewing a static image, we suggest that interacting with a robot can lead to a different and richer (sensorial and emotional) experience, especially for androids. Becker, Asano, Ogawa, Nishio, and Ishiguro [48], for instance, observed 24 people (seventeen Austrians, three Germans, three Swiss, and one British) interacting with a Geminoid HI-1 and noted that the majority of the comments (45 on 70 comments) that reported included some positive feelings. More studies are therefore needed to evaluate whether (and which) direct interactions could attenuate (or exacerbate) the perceived fear of damage to humans and their identity.

Another limitation concerns the fact that the participants of our studies were all Italians. This raises the question whether the present findings would generalize to other national samples and cultural contexts. For example, researchers [9] have suggested that, compared to Westerners, Japanese people

might be more positively disposed towards robots in general and androids in particular because East Asian culture is more tolerant toward objects that cross category boundaries. Having said that, previous empirical work provides no evidence for cultural differences between the West and East. Specifically, in a study with Japanese participants, Kamide et al. [4] found a pattern of results that is largely similar to our findings: compared to humanoids, androids were judged as more human-like and as more of a threat to humans and their identity. In addition, survey studies showed that Japanese and European respondents [50] did not differ substantially in their attitudes toward robots and in their belief whether robots should look like humans (see also Bartneck et al. [49] for a US and Japanese comparison). Nevertheless, we recommend that future studies should further explore potential cultural differences. It may also be of interest to examine how the human-machine divide is affected by other contextual effects relating to for example educational background or religious beliefs (see MacDorman and Entzari [12]).

Future research should also focus on gaining a better understanding of the type of threat that robots, and especially androids, pose. In our study we relied on the Kamide et al. scale [4] to assess the perceived damage to humans and their identity, as this scale has good psychometric properties and was created following rigorous piloting. That said, we acknowledge that this scale includes items assessing different fears than those relating specifically to threat to human identity (e.g., fear that humans could lose control, fear of being physically harmed, concerns about losing identity value and specificity, etc.). Even though we found in our studies that these different fears were highly correlated and that the pattern of results is similar for each of the items, future studies are needed to examine whether different types of robots pose different types of threat (e.g., androids might threaten human identity, whereas mechanical and humanoid robots may be more threatening in arousing fears that robots replace humans in the workplace). It may also be worthwhile to examine whether androids represent not only a threat to humans and human identity but also a threat to the natural world more generally. Finally, future research should focus on identifying ways to prevent this threat to human distinctiveness to arise. Studies in social psychology would suggest that increasing the differences between humans and robots would preserve the human need for distinctiveness even when facing robots high in anthropomorphic appearance. For instance, adding a distinctive marker on androids (e.g., a tattoo or a specific dress) would create a visible difference and this would facilitate the identification of these robots. Note however that this would not alter the fact that androids are mechanical agents with a biological appearance. According to recent studies [51] stimuli that merge human and non-human features elicit a state of discomfort and fear as they activate competing interpretations. Following this line of reasoning,

adding a marker may not be sufficient to preserve human distinctiveness, as the threatening element of androids would be the mix between human and mechanical features. Future studies should also investigate whether other robot features, beyond those relating to that of appearance, can contribute to overcoming the resistance towards this technology. For example, Sorbello, Chella, Giardina, Mishio, and Ishiguro [52], suggested that the robot's ability to show empathy towards humans would improve its acceptance (see also, [53,56]). Results of their study are fascinating and emotional reactions toward android are at odds with those of Gray and Wegner [27] showing that the ability of experiencing and understanding emotions increased rather than decreased robot Kaspar's (<http://www.herts.ac.uk/kaspar>) creepiness. One way to reconcile these contrasting findings is that people generally expect a match between the robot's appearance and behavior (see also [57]). Although the present research was not designed to address this issue, it provides some indirect evidence in support of this reasoning. We found that, compared to humanoids and mechanical robots, androids were judged as looking most like humans but also as behaving somehow more humanly, given that they were rated to possess to a greater extent qualities typical of human mind and nature. Interestingly, the higher attribution of human mind and human traits did not account for the higher threat to distinctiveness and perceived damage to humans and their identity elicited by robots with an anthropomorphic appearance. This finding leaves open the possibility that humanlike behavior in androids does not increase the negative feelings towards these robots. However, further studies are needed to further explore this possibility.

8 Conclusion

In the present research we showed that robots that look “too human” and can therefore be mistaken to be one of us give rise to concerns that their entering in the society would negatively impact on humans as a group. To avoid people resistance, roboticists should develop robots whose appearance does not challenge the psychological distinction between humans and mechanical agents.

Acknowledgements The research for this paper was financially supported by a doctorate grant awarded by the University of Trento to F. Ferrari. Portions of the data of Study 1 have been analyzed for a different purpose and presented in form of a proceeding at “Evaluating Social Robots”, The 13th International Conference on Intelligent Autonomous Systems, July 18, 2014, Padova, Italy.

Authors contribution Francesco Ferrari, Maria Paola Paladino and Jolanda Jetten developed the study concept. Francesco Ferrari and Maria Paola Paladino designed the studies. Francesco Ferrari prepared the experimental material, collected and analyzed the data. Francesco

Ferrari and Maria Paola Paladino drafted the manuscript. Jolanda Jetten edited and contributed to the critical revisions of the manuscript. All the authors read and approved the final version for submission.

References

- Kanda T, Ishiguro H, & Ishida T. 2001. Psychological analysis on human-robot interaction. In *Robotics and Automation, 2001. Proceedings 2001 ICRA. IEEE International Conference on*, Vol. 4, pp. 4166-4173. IEEE. DOI: 10.1109/ROBOT.2001.933269.
- Lee, K. M., Jung, Y., Kim, J., & Kim, S. R. (2006). Are physically embodied social agents better than disembodied social agents?: The effects of physical embodiment, tactile interaction, and people's loneliness in humanrobot interaction. *International Journal of Human-Computer Studies*, 64(10), 962-973. DOI: 10.1016/j.ijhcs.2006.05.002.
- European Commission, Special Eurobarometer 382, Public Attitudes Toward Robots (February-March, 2012). TNS Opinion & Social, Brussels [Producer]. http://ec.europa.eu/public_opinion/archives/ebs/ebs_382_en.pdf. Accessed 22th May 2015.
- Kamide, H., Mae, Y., Kawabe, K., Shigemi, S., & Arai, T. (2012, May). A psychological scale for general impressions of humanoids. In *Robotics and Automation (ICRA), 2012 IEEE International Conference on* (pp. 4030-4037). DOI: 10.1080/01691864.2013.751159.
- Mori, M. 1970. The uncanny valley. *Energy*, 7(4), 33-35.
- Mori M, MacDorman KF, & Kageki, N. 2012. "The Uncanny Valley [From the Field]", *Robotics & Automation Magazine*, IEEE, vol.19, no.2, pp.98,100, June 2012. DOI: 10.1109/MRA.2012.2192811.
- Ramey, C. H. (2005). The uncanny valley of similarities concerning abortion, baldness, heaps of sand, and humanlike robots. In *Proceedings of views of the uncanny valley workshop: IEEE-RAS international conference on humanoid robots* (pp. 8-13).
- Kaplan F. 2004. Who is Afraid of the Humanoid? Investigating Cultural Differences in the Acceptance of Robots. In *International Journal of Humanoid Robotics*, vol.1, no.3, pp.1-16. DOI: 10.1142/S0219843604000289.
- MacDorman, K. F., Vasudevan, S. K., & Ho, C. C. (2009). Does Japan really have robot mania? Comparing attitudes by implicit and explicit measures. *AI & society*, 23(4), 485-510. DOI: 10.1007/s00146-008-0181-2.
- MacDorman, K. F., & Ishiguro, H. (2006). The uncanny advantage of using androids in cognitive and social science research. *Interaction Studies*, 7(3), 297-337. DOI: 10.1075/is.7.3.03mac.
- Rosenthal-von der Pütten, A. M., Krmer, N. C., Becker-Asano, C., Ogawa, K., Nishio, S., & Ishiguro, H. (2014). The Uncanny in the Wild. Analysis of Unscripted Human-Android Interaction in the Field. *International Journal of Social Robotics*, 6(1), 67-83. DOI: 10.1007/s12369-013-0198-7.
- MacDorman, K. F., & Entezari, S. O. (2015). Individual differences predict sensitivity to the uncanny valley. *Interaction Studies*, 16(2), 141-172. DOI: 10.1075/is.16.2.01mac.
- Tajfel, H., & Turner, J. C. (1979). An integrative theory of intergroup conflict. *The social psychology of intergroup relations*, 33(47), 74. DOI: 10.1146/annurev.ps.33.020182.000245
- Brewer, M. B. (1991). The social self: On being the same and different at the same time. *Personality and social psychology bulletin*, 17(5), 475-482. DOI: 10.1177/0146167291175001.
- Jetten, J., Spears, R., & Manstead, A. S. (1996). Intergroup norms and intergroup discrimination: distinctive self-categorization and social identity effects. *Journal of personality and social psychology*, 71(6), 1222. DOI: 10.1037/0022-3514.71.6.1222.
- Jetten, J., Spears, R., & Manstead, A. S. (1997). Distinctiveness threat and prototypicality: Combined effects on intergroup discrimination and collective self-esteem. *European Journal of Social Psychology*, 27(6), 635-657. DOI: 10.1002/(SICI)1099-0992(199711/12)27:6<635::CO;2-#>
- Haslam N. 2006. Dehumanization: An Integrative Review. In *Personality and Social Psychology Review*, vol.10, no. 3, pp. 252-264. DOI: 10.1207/s15327957pspr1003.4.
- Vaes, J., Leyens, J. P., Paola Paladino, M., & Pires Miranda, M. (2012). We are human, they are not: Driving forces behind outgroup dehumanisation and the humanisation of the in-group. *European Review of Social Psychology*, 23(1), 64-106. DOI:10.1080/10463283.2012.665250.
- Enz, S., Diruf, M., Spielhagen, C., Zoll, C., & Vargas, P. A. (2011). The social role of robots in the futureexplorative measurement of hopes and fears. *International Journal of Social Robotics*, 3(3), 263-271. DOI: 10.1007/s12369-011-0094-y.
- Hegel, F., Eyssel, F., & Wrede, B. (2010, September). The social robot 'flobi': Key concepts of industrial design. In *RO-MAN, 2010 IEEE* (pp. 107-112). IEEE. DOI: 10.1109/ROMAN.2010.5598691.
- Ishiguro, H., Ono, T., Imai, M., Maeda, T., Kanda, T., & Nakatsu, R. (2001). Robovie: an interactive humanoid robot. *Industrial robot: An international journal*,28(6), 498-504. DOI: 10.1108/01439910110410051.
- Zecca, M., Mizoguchi, Y., Endo, K., Iida, F., Kawabata, Y., Endo, N., Itoh, K., & Takanishi, A. (2009, September). Whole body emotion expressions for KOBIAN humanoid robotpreliminary experiments with different Emotional patterns. In *Robot and Human Interactive Communication, 2009. RO-MAN 2009. The 18th IEEE International Symposium on* (pp. 381-386). IEEE. DOI: 10.1109/RO-MAN.2009.5326184.
- Hornsey, M.J., & Jetten, J. (2003). Not being what you claim to be: Impostors as sources of group threat. *European Journal of Social Psychology*, 33, 639-657. DOI: 10.1002/ejsp.176.
- Jetten, J., Summerville, N., Hornsey, M.J., & Mewse, A. J. (2005). When differences matter: Intergroup distinctiveness and the evaluation of impostors. *European Journal of Social Psychology*, 35, 609-620. DOI: 10.1002/ejsp.282.
- Warner, R., Hornsey, M. J., & Jetten, J. (2007). Why minority group members resent impostors. *European Journal of Social Psychology*, 37(1), 1-17. DOI: 10.1002/ejsp.332.
- Jetten, J., & Hornsey, M. J. (Eds.). (2010). *Rebels in groups: Disent, deviance, difference, and defiance*. John Wiley & Sons., pp. 174. DOI: 10.1002/ejsp.332.
- Gray K, & Wegner DM. 2012. Feeling robots and human zombies: Mind perception and the uncanny valley. *Cognition*, 125(1), 125130. DOI: 10.1016/j.cognition.2012.06.007.
- Ferrari F., & Paladino M. P. Validation of the Psychological Scale of General Impressions of Humanoids in an Italian Sample. *Workshop Proceedings of IAS-13, 13th International Conference on Intelligent Autonomous Systems, Padova (Italy) July 15-19,2014*, ISBN: 978-88-95872-06-3, pp.436-441.
- Gray HM, Gray K, & Wegner DM. 2007. Dimensions of Mind Perception. In *Science*, vol. 315, no. 5812 p. 619. (with supportive materials). DOI: 10.1126/science.1134475.
- <http://www.afhayes.com/spss-sas-and-mplus-macros-and-code.html>
- Hahn-Holbrook, J., Holt-Lunstad, J., Holbrook, C., Coyne, S. M., & Lawson, E. T. (2011). Maternal defense: Breast feeding increases aggression by reducing stress. *Psychological Science*, 22, 1288-1295. DOI: 10.1177/0956797611420729.
- Legault, L., Gutsell, J. N., & Inzlicht, M. (2011). Ironic effects of antiprejudice messages: How motivational interventions can reduce (but also increase) prejudice. *Psychological Science*, 22, 1472-1477. DOI: 10.1177/0956797611427918.

33. Waytz, A., Heafner, J., & Epley, N. (2014). The mind in the machine: Anthropomorphism increases trust in an autonomous vehicle. *Journal of Experimental Social Psychology*, 52, 113-117. DOI: 10.1016/j.jesp.2014.01.005.
34. Fritz MS, MacKinnon DP. Required sample size to detect the mediated effect. *Psychological Science*. 2007;18:233239. DOI: 10.1111/j.1467-9280.2007.01882.x.
35. Preacher KJ, and Hayes AF. 2008. Asymptotic and resampling strategies for assessing and comparing indirect effects in multiple mediator models. *Behavior research methods*, 40(3), 879-891. DOI: 10.3758/BRM.40.3.879.
36. Mazzei, D., Billeci, L., Armato, A., Lazzeri, N., Cisternino, A., Pioggia, G., Iglizoi, R., Muratori, F., Ahluwalia, A., De Rossi, D. (2010). The FACE of autism. *Proceedings - IEEE International Workshop on Robot and Human Interactive Communication*, art. no. 5598683, pp. 791-796. DOI: 10.1109/ROMAN.2010.5598683.
37. Mazzei, D., Lazzeri, N., Billeci, L., Iglizoi, R., Mancini, A., Ahluwalia, A., Muratori, F., De Rossi, D. (2011). Development and evaluation of a social robot platform for therapy in autism. *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS*, art. no. 6091119, pp. 4515-4518. DOI: 10.1109/IEMBS.2011.6091119.
38. Luke, M. A., & Maio, G. R. (2009). Oh the humanity! Humanity-esteem and its social importance. *Journal of Research in Personality*, 43(4), 586-601. DOI: 10.1016/j.jrp.2009.03.001.
39. Hyman HH. (1955). *Survey Design and Analysis: Principles, Cases, and Procedures*. Glencoe, IL: Free Press. (pp.280). DOI: 10.1177/001316445601600312.
40. Judd, C. M., & Kenny, D. A. (1981). Process analysis estimating mediation in treatment evaluations. *Evaluation review*, 5(5), 602-619. DOI: 10.1177/0193841X8100500502.
41. Baron, R. M., & Kenny, D. A. (1986). The moderator-mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of personality and social psychology*, 51(6), 1173. DOI: 10.1037/0022-3514.51.6.1173.
42. Duffy BR. (2003). Anthropomorphism and the social robot, 42, 177190. doi:10.1016/S0921-8890(02)00374-3. DOI: 10.1016/S0921-8890(02)00374-3.
43. Fink, J. (2012). Anthropomorphism and Human Likeness in the Design of Robots and Human-Robot Interaction, 199208. DOI: 10.1007/978-3-642-34103-8_20.
44. Mitchell, W. J., Ho, C. C., Patel, H., & MacDorman, K. F. (2011). Does social desirability bias favor humans? Explicit/implicit evaluations of synthesized speech support a new HCI model of impression management. *Computers in Human Behavior*, 27(1), 402-412. DOI: 10.1016/j.chb.2010.09.002.
45. MacDorman, K. F., Coram, J. A., Ho, C. C., & Patel, H. (2010). Gender differences in the impact of presentational factors in human character animation on decisions in ethical dilemmas. *Presence: Teleoperators and Virtual Environments*, 19(3), 213-229. DOI: 10.1162/pres.19.3.213.
46. Heerink, M., Krse, B., Evers, V., & Wielinga, B. (2010). Assessing acceptance of assistive social agent technology by older adults: the almere model. *International journal of social robotics*, 2(4), 361-375. DOI: 10.1007/s12369-010-0068-5.
47. Rosenthal-von der Pten, A. M., & Krmer, N. C. (2014). How design characteristics of robots determine evaluation and uncanny valley related responses. *Computers in Human Behavior*, 36, 422-439. DOI: 10.1016/j.chb.2014.03.066.
48. Becker-Asano, C., Ogawa, K., Nishio, S., & Ishiguro, H. (2010). Exploring the uncanny valley with Geminoid HI-1 in a real-world application. In *Proceedings of IADIS International Conference Interfaces and Human Computer Interaction* (pp. 121-128). ISBN: 978-972-8939-18-2.
49. Bartneck, C. (2008, August). Who like androids more: Japanese or US Americans?. In *Robot and Human Interactive Communication*, 2008. RO-MAN 2008. The 17th IEEE International Symposium on (pp. 553-557). IEEE. DOI: 10.1109/ROMAN.2008.4600724.
50. Haring, K. S., Mougnot, C., Ono, F., & Watanabe, K. (2014). Cultural differences in perception and attitude towards robots. *International Journal of Affective Engineering*, 13(3), 149-157. DOI: 10.5057/ijae.13.149.
51. Burleigh, T. J., Schoenherr, J. R., & Lacroix, G. L. (2013). Does the uncanny valley exist? An empirical test of the relationship between eeriness and the human likeness of digitally created faces. *Computers in Human Behavior*, 29(3), 759-771. DOI: 10.1162/pres.16.4.337.
52. Sorbello, R., Chella, A., Giardina, M., Nishio, S., & Ishiguro, H. (2014, July). An Architecture for Telenoid Robot as Empathic Conversational Android Companion for Elderly People. In the 13th International Conference on Intelligent Autonomous Systems (IAS-13), Padova, Italy.
53. Damiano, L., Dumouchel, P., & Lehmann, H. (2014). Towards HumanRobot Affective Co-evolution Overcoming Oppositions in Constructing Emotions and Empathy. *International Journal of Social Robotics*, 7(1), 7-18.
54. Leite, I., Castellano, G., Pereira, A., Martinho, C., & Paiva, A. (2014). Empathic robots for long-term interaction. *International Journal of Social Robotics*, 6(3), 329-341
55. Asada, M. (2014). Towards artificial empathy. *International Journal of Social Robotics*, 7(1), 19-33.
56. Lim, A., & Okuno, H. G. (2014). A Recipe for Empathy. *International Journal of Social Robotics*, 7(1), 35-49.
57. Saygin, A.P., Chaminade, T., Ishiguro, H., Driver, J. & Frith, C. (2012) The thing that should not be: Predictive coding and the uncanny valley in perceiving human and humanoid robot actions. *Social Cognitive Affective Neuroscience*, 7(4):413-22; PMID: 21515639. DOI: 10.1093/scan/nsr025.

Biographies

FRANCESCO FERRARI is a PhD candidate of the Department of Psychology and Cognitive Science of University of Trento, Italy. His research interests concern social and cognitive processes involved when observing/interacting with social robots. In his PhD thesis he focused on social robots anthropomorphic appearance and peoples concerns about potential negative impacts that this technology.

MARIA PAOLA PALADINO is a Professor of Social Psychology at the University of Trento, Italy. Her research interests are how people perceive, relate and behave toward other people or social relevant entities (i.e. robots). She was one of the original proponents of the infra-humanization theory, whose approach and hypothesis has strongly influenced the current knowledge on people dehumanization and object anthropomorphism.

JOLANDA JETTEN is a Professor of Social Psychology at the University of Queensland. Her research is concerned with group processes, social identity and intergroup relations, as represented by her most recent books: *Rebels in groups: Dissent, deviance, difference, and defiance* (coedited with Matthew Hornsey; 2011) and *The social cure: Identity, health and well-being* (coedited with Catherine Haslam and Alex Haslam; 2012).