Strategies for the Promotion of Humanity Attribution to Outgroups

Loris Vezzali¹, Dora Capozza², Gian Antonio Di Bernardo¹,
Rossella Falvo², Sofia Stathi³, Gordon Hodson⁴

¹University of Modena and Reggio Emilia, Italy
²University of Padova, Italy
³University of Greenwich, UK
⁴Brock University, Canada

Correspondence concerning this article should be addressed to Loris Vezzali, University of Modena and Reggio Emilia, viale Allegri 9, 42121, Reggio Emilia, Italy. E-mail: loris.vezzali@unimore.it
Tel: + 39 0522 523006

Conflict of interest disclosure

The authors declare no conflict of interest.

Word count: 11,881 words (text)
Strategies for the Promotion of Humanity Attribution to Outgroups

Abstract

Outgroup dehumanisation, the denial of full humanity to outgroups relative to ingroups, is pervasive in many contemporary societies. The aim of the present work is to review effective strategies aimed at fostering outgroup humanity attribution. After presenting the main models of humanity attribution, we differentiate two types of strategies. Outgroup-specific strategies are focused on a target outgroup, therefore their effectiveness is more dependent upon the specific intergroup relationship. These include intergroup contact, meta-humanisation, and social categorisation. In contrast, outgroup-independent strategies are not inherently linked to a target outgroup, implying that their effectiveness is less dependent on the specific intergroup dynamics under consideration. These involve human-animal similarity and secure relationship attachment. We provide evidence for the effectiveness of these strategies and their underlying processes, showcasing our research program within the larger literature. In so doing, we take into account the distinction between blatant and subtle dehumanisation, and conclude with suggestions for future research.

Keywords: dehumanisation, humanity attribution, intergroup relations, prejudice, interventions.
One of the key aims of classic and contemporary psychosocial research has been to investigate different forms of prejudice and discrimination. Outgroup dehumanisation, that for the present purposes we define as the denial of full humanity to outgroups (vs. ingroups), and for simplicity we refer to as ‘dehumanisation’, represents a particular expression of negativity that has detrimental consequences, such as justifying a wide range of atrocities that target the outgroup (e.g., Haslam, 2021). Initially, researchers focused on blatant forms of dehumanisation, such as the explicit denial of humanity to outgroup members, psychologically equating them to animals and parasites, in an effort to explain transgressions during World War II (Kelman, 1976). Building on this body of work, research subsequently focused to a greater extent on more subtle forms of denial of humanity to the outgroup. This line of research has culminated in several reviews of outgroup humanity attribution¹ (see Haslam, 2006; Haslam & Loughnan, 2014; Haslam & Stratemeyer, 2016; Leyens et al., 2007; Vaes et al., 2012). Whereas past reviews have documented the negative aspects of the denial of humanity, our review focuses on solutions, specifically regarding strategies to humanise the outgroup. We identify five main strategies: (1) intergroup contact; (2) meta-humanisation; (3) social categorisation; (4) human-animal divide perceptions; and (5) secure attachment. Further, we group them into two types of strategies: outgroup-specific and outgroup-independent strategies. In presenting each strategy we will note the type – blatant or subtle – of humanity attribution under investigation.

Theoretical Approaches and Evidence of Humanity Attribution

Significant scholarly attention has been devoted to processes that help to explain and attenuate the widespread denial of humanity to outgroup members, as well as to enhancing humanity attribution. One important aspect concerns the distinction between blatant and subtle forms of dehumanisation. Blatant forms of humanity attribution refer to the explicit and unambiguous denial of humanity of another group, which is typical in highly conflictual contexts such as wars or open conflict. Unsurprisingly, this form can lead to a complete disregard of life and extreme forms of violence (Bandura et al., 1975). A horrifying example is provided by the 1994
Rwandan genocide, where Tutsis were characterised by Hutus as “cockroaches”, were butchered en masse with machetes, and living babies were removed from wombs (note that violence also occurred toward moderate or less-clearly dehumanised Hutus) (Human Rights Watch, 1996). Bandura (1990; see also Opotow, 1990; Staub, 1987) focused on the role of moral disengagement to explain the causes underlying such behaviour: by excluding outgroup members from the human category, individuals’ inhibitions to engage in extreme actions are removed. Although strong forms of dehumanisation can be characterised as a remnant of the past and/or associated with extreme intergroup conflict, evidence of its occurrence is still present in daily life. For instance, cases where Black athletes or politicians are associated with apes are sadly common (e.g., former U.S. President Barack Obama depicted as a chimpanzee by a political opponent, EEWMagazine, 2011).

Kteily and Bruneau (2017) argued for the need to also investigate overt modes to deny humanity to the outgroup. Kteily et al. (2015) introduced a blatant dehumanisation scale, taking advantage of the popular “Ascent of Humans” diagram. Participants are asked to position the target group on such a diagram by using a response scale anchored to 0 (ape-like human ancestors) and 100 (advanced modern humans). In seven studies the authors demonstrated the validity of this measure, which predicted prejudice toward several outgroups over and above subtle measures of humanity attribution. Bruneau et al. (2018) provided convincing evidence with samples from four European countries that Muslim refugees are blatantly dehumanised, and that dehumanisation is associated with anti-refugee behaviour over and above the effects of several variables, including prejudice.

This current surge of studies examining dehumanisation as a blatant form of humanity attribution is, in part, a reaction to the more pervasive study of subtle forms of denial of humanity of the outgroup during the first two decades of the 21st century. As one example, infrahumanisation theory draws on the concept of human essence, consisting of what is psychologically considered common among social groups, and ethnocentrism, the notion that the ingroup is preferred over the outgroup. Starting from these premises and by relying on social identity theories of intergroup
relations (Tajfel & Turner, 1979), infrahumanisation theory predicts that human essence is assigned more to ingroup than to outgroup members (Leyens et al., 2007). In its initial and more popular operationalisation, infrahumanisation is based on the perceived distinction between uniquely human (i.e., secondary) and non-uniquely (i.e., primary) human emotions. Whereas uniquely human emotions (e.g., remorse, hope) are widely thought to be experienced by humans, non-uniquely human emotions (e.g., anger, joy) are widely thought to be experienced by both humans and animals. There is now consistent evidence that individuals ascribe uniquely human emotions more to the ingroup than to the outgroup, while non-uniquely human emotions are equivalently ascribed to ingroup and outgroup members (for reviews, see Haslam & Loughnan, 2014; Leyens et al., 2007; McLoughlin & Over, 2018; Vaes et al., 2012). Of note, research also moved beyond mere emotion attribution, demonstrating that infrahumanisation may also occur when considering the distinction between non-uniquely and uniquely human traits (e.g., Capozza, Trifiletti, et al., 2013; Costello & Hodson, 2010, 2014a; Haslam, 2006; Hodson & Costello, 2007).

The dual model of dehumanisation proposed by Haslam (Haslam, 2006; Haslam et al., 2008) also focuses on subtle forms of humanity attribution. According to these authors, humans can be likened to animals, but also to other entities, such as inanimate objects, including robots or machines. Haslam et al. identified two dimensions along which to define humans. The first is human uniqueness (HU), which differentiates human from animals, thus tapping the animalistic form of dehumanisation. Items measuring human uniqueness refer, for instance, to rationality or morality. In contrast, human nature (HN) relates to human essence and serves to assess mechanistic dehumanisation, differentiating humans from robots or machines. Human nature is generally measured with traits related to warmth and emotionality. There is now ample evidence supporting the basic premises of this theoretical framework (for reviews, see Haslam & Loughnan, 2014; Vaes et al., 2012).

Research has shown that subtle forms of dehumanisation are associated with a wide range of detrimental outcomes, such as reduced outgroup approach intentions (Capozza et al., 2016), lower
perspective-taking and empathy toward the outgroup (Čehajić et al., 2009), increased prejudice (Costello & Hodson, 2010, 2011, 2014b; Hodson & Costello, 2007), enhanced tendencies for intergroup aggression (Greitemeyer & McLatchie, 2011), reduced intergroup forgiveness (Stathi et al., 2017, Study 1). Despite the detrimental consequences of blatant and subtle forms of dehumanisation, comparatively few of studies have focused on how to promote the attribution of humanity to the outgroup (cf. Haslam & Loughnan, 2014). Yet there has accrued, in recent years, a sufficient body of evidence to warrant a review.

Two Types of Strategies to Foster Humanity Attribution

We identify five strategies, conceptually differentiating them into two main categories. The first includes outgroup-specific strategies, specifically intergroup contact, meta-humanisation, and social categorisation. These strategies are mainly determined by the specific intergroup dynamic. Contact with an outgroup can shape how this specific outgroup is appraised and evaluated, it may determine emotional responses and cognitive processes associated with it (like psychological distance or threat), and the resulting humanity attribution. Similarly, meta-humanisation, that is the perception that one’s ingroup is perceived as possessing dignified, human-like qualities by an outgroup, presumably changes as a function of the outgroup taken into consideration. According to Vorauer et al. (1998), meta-perceptions are a relational concept, and thus determined by the extent to which the outgroup is expected to attribute humanity to the self and to the perceiver’s ingroup. For example, believing that outgroup members attribute more humanity to their ingroup than to outgroup members makes this meta-perception a relevant intergroup construct differentiating the two groups. With respect to social categorisation, in line with self-categorisation theory (J. C. Turner et al., 1987), perceptions of ingroup similarity and outgroup dissimilarity are inextricably linked in determining social (re)categorisation processes. According to the meta-contrast principle, the nature of the intergroup context is comparative, with group perceptions depending on the salient groups. Specifically, social categorisation depends on perceived relative similarities and differences between groups. Categorising ingroup and outgroup as a single group (common ingroup
categorisation) may be less effective if the two groups are perceived as very different in their defining characteristics. As an example, perceptions of differences between the two groups may lower perceptions (and therefore strength) of the common identity.

Therefore, strategies based on social categorisation strictly depend upon the specific intergroup relationship within which they occur. We are not, however, claiming that the effects of outgroup-specific strategies are confined to the specific intergroup relationship. Rather, outgroup-specific strategies are primarily defined by the relationship with the specific outgroup for which they are tailored (e.g., contact is intended with a specific outgroup; meta-humanisation refers to humanizing perceptions from a specific outgroup). However, the fact that these strategies refer to a specific outgroup does not eliminate the possibility that some generalization or “spillover” effects can occur. For instance, it is now established that contact with a specific outgroup can change attitudes toward outgroups uninvolved in the contact situation (Pettigrew, 2009; Vezzali et al., 2021).

*Outgroup-independent strategies* include the strategies based on the human-animal divide and attachment orientations. Concerning the human-animal divide, fostering perceptions that humans and animals are similar is not linked to a specific human outgroup in question, but represents a more general strategy tapping on similarities between humans in contrast to animals. Therefore, its effectiveness is less likely to depend on the specific social outgroup under investigation (i.e., the strategy in itself is not focused on any human outgroup). Similarly, strategies based on priming the secure attachment orientation are relatively independent of the specific intergroup relationship. Rather they are based on priming concepts related to the individual as a person and his/her attachment figures or interpersonal relationships. Being relatively independent of the specific intergroup relation, these strategies may be effective in enhancing humanity attribution toward a wide variety of outgroups simultaneously.

*Outgroup-Specific Strategies*

*Intergroup contact*
Direct, face-to-face intergroup contact represents one of the most widely endorsed and effective strategies for prejudice reduction (Hodson & Hewstone, 2013; Pettigrew & Tropp, 2006; Vezzali & Stathi, 2017, 2021). That is, contact between representatives of their respective groups can develop more positive attitudes toward the other’s group as a whole. There is now also evidence that indirect forms of contact, that is contact that is not experienced face-to-face, are effective in attenuating prejudice. Indirect contact forms that have received substantial attention from research, and are related to humanity attribution, are extended and vicarious contact, which are based on the notion that knowing that ingroup members have outgroup friends (or watching intergroup contact vicariously) improves outgroup attitudes (Vezzali et al., 2014; Wright et al., 1997; Zhou et al., 2019). A further indirect contact strategy is imagined contact, the mental simulation of a positive encounter with a member of the outgroup (Crisp et al., 2010; Crisp & Turner, 2012; Hodson et al., 2015; see also White et al., 2020).

Until recently research has largely overlooked the effects of contact on humanity attribution. There is now, however, substantial support for the effectiveness of contact in this domain (Capozza, Falvo, Di Bernardo, et al., 2014; Vezzali & Stathi, 2021, Chapter 5). An experimental demonstration of the effects of contact has been provided by Capozza et al. (2017, Study 2). In this study, we used an approach-training technique (see Kawakami et al., 2007) to manipulate contact. Participants were 57 Italian university students rating Moroccans as the outgroup. The study included one experimental and two control conditions. In the experimental condition stimuli were represented by six outgroup faces and two geometrical figures (two ovals). Each stimulus was presented 12 times, with a manikin situated above the stimulus in six trials and below it in the remaining six trials. Participants were instructed to move the manikin (representing the self) on the screen toward outgroup faces (for each of the six outgroup faces, upwards in six trials, downwards in the other six trials). In so doing, they were also asked to imagine cooperative contact with the Moroccan person. The combination of approach movements and mental simulation of contact was meant to make the contact manipulation more impactful. Participants also had to simulate avoidance
of geometrical figures (the two ovals), moving the manikin away from them (for each oval, upwards in six trials, downwards in the other six trials). In one control condition, Moroccan faces were replaced by neutral stimuli (six images of furniture). The participants’ task was to move the manikin toward these stimuli and away from the two ovals. We included a further control condition (sideways-control condition), which used the same stimuli as the contact condition; we aimed to exclude that an eventual increase in humanity attribution was due to outgroup faces per se rather than repeated approach to outgroup members. Participants in this condition moved the manikin toward the right in response to outgroup faces, and toward the left in response to the two ovals (see Figure 1). Humanity attribution was assessed via the ascription of four uniquely (e.g., morality) and four non-uniquely (e.g., impulsiveness) human traits, rated on a 7-point scale ranging from 1 to 7. We found that uniquely human traits were attributed to a greater extent to the outgroup in the contact condition than in the two control conditions \(M_{\text{contact}} = 4.74, M_{\text{control}} = 3.64, M_{\text{sideways control}} = 4.06, ds = 1.39\) and \(.91\) for the first and second comparisons, respectively); no difference emerged for non-uniquely human traits across conditions.

Falvo et al. (2014) provided experimental longitudinal evidence by utilising the imagined contact paradigm. In this case, we focused on individuals with intellectual disability as a stigmatised minority. Participants were 164 adults without disability. Participants assigned to the experimental condition were asked to imagine a positive and pleasant encounter with a person with intellectual disability. They were further asked to keep their eyes closed while imagining a detailed encounter from a third person perspective, techniques that strengthen the effects of imagined contact (Crisp & Turner, 2012). In the control condition, participants imagined an outdoor landscape. Participants were then administered a questionnaire (T1), where they were asked to assign six uniquely human (three positive, e.g., hope, and three negative, e.g., resentment) and six non-uniquely (three positive, e.g. pleasure, and three negative, e.g., sadness) human emotions to the outgroup. An index of humanity perceptions was computed by calculating the difference between the number of non-
uniquely and uniquely human emotions, with higher scores representing greater ascription of non-uniquely than uniquely emotions and thus lower humanity attribution. Participants were contacted one month later to complete a similar questionnaire (T2). Results revealed that the denial of humanity was higher in the control than in the experimental condition ($M_s = 1.28$ vs. $.77$, $d = .33$), with effects persisting over time ($M_s = 1.21$ vs. $.84$, $d = .24$).

In a cross-sectional study Stathi et al. (2017, Study 2) extended these results by testing the association between contact and humanity attribution within a conflictual context, and with a measure of blatant dehumanisation. The study was conducted in Cyprus considering the historically violent relation between Greek and Turkish Cypriots where dehumanisation could indeed manifest. We further explored the construct of intergroup forgiveness as a requisite step toward the reconciliation process. We recruited a community sample of 86 Turkish Cypriots (rating Greek Cypriots). Dehumanisation was assessed by asking participants to indicate the degree to which each of eight traits such as ‘beast’, ‘mongrel’, ‘human’ (reversed) can describe Greek Cypriots. Forgiveness was assessed with two items: “I think that Turkish Cypriots should forgive Greek Cypriots misdeeds”; “Cyprus will never move forward until Turkish Cypriots forgive Greek Cypriots” (reverse-scored). The contact index was computed by multiplying measures of quantity and quality of contact, with higher scores reflecting frequent positive contact. Results revealed a negative correlation between contact and dehumanisation ($r = -.32$). A path model revealed that contact was negatively associated with dehumanisation ($\beta = -.24$), which in turn was negatively associated with forgiveness ($\beta = -.28$); the indirect effect was significant (95% bootstrap CI comprised between .001 and .029).

Further ecological validity for the role of contact was provided by an experimental intervention in the field conducted with a child sample (Vezzali et al., 2012). This study aimed at testing whether imagined contact was effective when tested in the field with a multi-session intervention, and when considering psychological constructs arguably difficult to change such as humanity attribution. Participants were 34 Italian fourth-graders, with immigrant children serving as
the outgroup. Participants took part in a 3-week contact intervention, with each session lasting approximately 30 minutes. Each week participants were divided in small groups of five-six children (with a researcher guiding the sessions). In their small group, participants were asked to imagine having a positive interaction with an immigrant child. In order to avoid subtyping of the imagined contact partner, each week we systematically varied the setting of the imagined encounter (at school, in the neighbourhoods, at the park) and the outgroup member imagined. We took multiple steps to reinforce the manipulation. First, participants were asked to imagine contact in detail, focusing for instance on what they said to become friends. Second, participants were given 10 minutes to write down what they imagined. Finally, they engaged in a discussion with peers from their group centred on what they had imagined. A questionnaire was administered one week after the last session. Children were asked to rate outgroup members by using four uniquely human (two positive, e.g., hope, and two negative, e.g., shame) and four non-uniquely human emotions (two positive, e.g., excitement, and two negative, e.g., fear). As expected, infrahumanisation emerged in the control condition, where participants assigned more non-uniquely than uniquely human emotions to outgroup members ($M_s = 4.34$ vs. 4.16, $d = .31$). In contrast, infrahumanisation did not emerge in the experimental condition, where we found no difference between non-uniquely and uniquely human emotions ($M_s = 4.46$ vs. 4.37, $d = .19$).

Overall, research demonstrates that contact is associated with greater attribution of humanity to a variety of outgroups, considering intergroup relationships based on religion (Bruneau et al., 2020; Tam et al., 2007, Study 1), nationality (Capozza et al., 2017; Rodriguez-Perez et al., 2011), ethnicity (Pinel et al., 2017; Prati & Loughnan, 2018, Studies 1 and 2; Stathi et al., 2017, Study 2), immigrant status (Capozza, Trifiletti, et al., 2013, Study 1; Visintin et al., 2017), geographical origin (Capozza, Falvo, et al., 2013; Capozza, Trifiletti, et al., 2013, Study 2), deviance from social norms (homeless: Falvo et al., 2015; sex offenders, testing humanity attribution as predictors of support for their rehabilitation: Viki et al., 2012, Study 4), disability (intellectual disability: Falvo et al., 2014), age (elderly: Drury et al., 2017), sexual orientation (Capozza, Falvo, Trifiletti, et al.,
The effects are also evident when considering extended contact (Andrighetto et al., 2012; Capozza, Falvo, Trifiletti, et al., 2014) and imagined contact (Falvo et al., 2014; Vezzali et al., 2012). The relationship between contact and (de)humanisation has emerged with correlational (Drury et al., 2017), longitudinal (Brown et al., 2007), and experimental (Capozza et al., 2017) methodologies. Importantly, evidence was also obtained in conflictual contexts (Cyprus, see Stathi et al., 2017, Study 2; Kosovo, see Andrighetto et al., 2012; Northern Ireland, see Tam et al., 2007, Study 1). Although few studies considered blatant dehumanisation, the findings from such studies are consistent in showing the beneficial effects of contact on humanity representations (see Bruneau et al., 2020; Stathi et al., 2017, Study 2; Viki et al., 2012, Study 4).

**Mediators.** In line with the dual nature of processes underlying direct (Pettigrew & Tropp, 2008), extended (Birtel et al., 2017; Vezzali et al., 2014), and imagined contact (Crisp et al., 2010; Vezzali et al., 2013), research reveals mediation effects of contact on humanity attribution via both affective (intergroup anxiety, empathy, trust) and cognitive (cognitive representations of groups, inclusion of the other in the self, ingroup and outgroup norms) factors. These findings are consistent with Haslam’s (2006) notion that dehumanisation is itself a function of both cognitive and affective factors. Below, we present evidence for these processes.

Focusing on affective variables, research has provided evidence for two principal mediators of contact research, intergroup anxiety and empathy. Capozza, Trifiletti, et al. (2013, Study 2) conducted a (correlational) on the relations between Northern and Southern Italians. Participants were 250 Northern Italian university students. Humanity attributions were assessed by asking participants to rate outgroup members on the possession of four uniquely (e.g., rationality) human traits. A path model revealed that quality of contact (i.e., positive/favourable) was associated with reduced intergroup anxiety ($\beta = -.30$) and with increased intergroup empathy ($\beta = .21$). In turn, intergroup anxiety was associated with lower ($\beta = -.21$), and intergroup empathy with higher humanity attribution ($\beta = .26$). Both indirect effects, showing that contact was associated with increased humanity attribution via lower intergroup anxiety and higher intergroup empathy, were
simultaneously significant (95% bootstrap CIs comprised between .021 and .218 for intergroup anxiety, and between .011 and .196 for intergroup empathy). These results were replicated cross-sectionally by Capozza, Falvo, et al. (2013): extended contact was also associated with greater humanity attribution via reduced anxiety and increased empathy. Research also supports the mediating role of outgroup trust, a further mediator of contact (Paterson et al., 2019). Evidence for mediation via trust was provided by the experimental studies of Capozza et al. (2017, Study 2) and Vezzali et al. (2012) described above in this section (see Figure 2 for an example). Outgroup trust also emerged as a significant mediator in the study by Capozza, Falvo, et al. (2013) considering the relation between Northerners and Southerners in Italy.

Figure 2 approximately here

With respect to cognitive mediators, we examined group representations. In line with the common ingroup identity model (Gaertner & Dovidio, 2000) and with the idea that outgroup humanity is preferentially attributed to ingroup (vs. outgroup) members (Haslam, 2006; Leyens et al., 2007), we reasoned that one-group perceptions should direct humanity attribution to former outgroup members who are subsequently included in the superordinate ingroup. These hypotheses were supported in two correlational studies, considering two different intergroup contexts (Italians vs. immigrants, and Northern vs. Southern Italians) from the perspective of the higher-status group (Capozza, Trifiletti, et al., 2013). In both studies, the effects of quality of contact on humanity attribution (of uniquely human traits) were mediated by stronger perceptions of belonging to a common group and lower perceptions to belong to two distinct groups.

We also examined three more cognitive variables that typically explain the effects of extended contact. According to Wright et al. (1997), knowing about positive interactions between ingroup and outgroup members should reduce the psychological distance toward outgroup members (therefore increasing the inclusion of the other in the self, IOS; Aron et al., 1992), and inform that ingroup and outgroup members are favourable to contact (therefore fostering perceptions of pro-contact ingroup and outgroup norms). We conducted two correlational studies considering the
relationship between Northerners and Southerners ($n = 254$ Northern Italians; Capozza, Falvo, et al., 2013), and heterosexuals and gay people ($n = 202$ heterosexual university students; Capozza, Falvo, Trifiletti, et al., 2014). Results revealed that extended contact was associated with stronger perceptions that ingroup (Capozza, Falvo, et al., 2013) and outgroup norms (Capozza, Falvo, Trifiletti, et al., 2014) support contact, and with greater IOS (Capozza, Falvo, Trifiletti, et al., 2014); in turn, IOS, ingroup and outgroup norms were positively associated with greater humanity attribution (using measures of trait attribution). Direct contact was indirectly associated with greater humanity attribution via increased IOS (Capozza, Falvo, et al., 2013), but not via ingroup or outgroup norms (Capozza, Falvo, et al., 2013; Capozza, Falvo, Trifiletti, et al., 2014).

In sum, the studies reviewed demonstrate the effectiveness of contact in fostering humanity attribution. It is worth noting, however, that research on contact and humanity attribution has important limitations. First, and understandably, research has only focused on high-status groups. Second, most research has been correlational, whereas experimental and/or longitudinal evidence (especially in naturalistic contexts) is needed.

Meta-humanisation

The social psychological literature on meta-perceptions, that is how people perceive to be seen and/or judged by others, has largely focused on interpersonal relations (Frey & Tropp, 2006), however more recently scholarly interest on intergroup meta-perceptions has witnessed an uptick (e.g., Landry et al., 2021; MacInnis & Hodson, 2012, 2013; Moore-Berg et al., 2020; Stathi et al., 2020; Vezzali, 2017; Vorauer, 2013). Meta-perceptions are informed by intuition and social cues, attention to people's verbal and nonverbal indications, previous experiences, projection of self-perception, and outgroup stereotypes (Ames, 2004; Frey & Tropp, 2006; Yzerbyt et al., 2009; Vezzali, 2017), which can lead to avoidance or hostility and conflict (Mendez et al., 2007; Shelton et al., 2006). When seeking to understand how humanity attribution can be enhanced, we thus need to acknowledge the emerging literature that focuses on meta-perceptions as instigators of (de)humanisation.
There is preliminary but strong evidence that perceiving that an outgroup sees one’s ingroup as human decreases dehumanisation of said outgroup, a process that is in turn associated with reduced prejudice and hostility (Kteily et al., 2016; Pavetich & Stathi, 2021a).

Introducing the concept of intergroup meta-dehumanisation, Kteily et al. (2016; see also Pavetich & Stathi, 2021a) found a reciprocal relationship between meta-dehumanisation and intergroup hostility, a path mediated by outgroup dehumanisation. Across 10 studies and more than 3,000 participants, the authors obtained evidence from various intergroup contexts, such as Palestinians and Israelis, Roma people and non-Roma Hungarians, Americans and Arabs, and demonstrated that meta-dehumanisation predicted outgroup hostility above and beyond meta-prejudice. This research highlighted that meta-dehumanisation creates cycles of reciprocal dehumanisation and intergroup aggression (see also Kteily & Bruneau 2017).

With this in mind, a new line of research examines the other end of the meta-(de)humanisation continuum: meta-humanisation. Research on meta-humanisation is still scarce but there is now evidence to suggest that perceiving that the ingroup is humanised by an outgroup can break the vicious cycle of dehumanisation (Kteily et al., 2016). For example, Kteily et al. found that priming meta-humanisation decreased reciprocal dehumanisation and prejudice toward the outgroup (Study 6) and intergroup threat (Study 7). Pavetich and Stathi (2021b) conducted seven experimental studies in the context of Muslim – non-Muslim relations, finding that meta-humanisation can break the cycle of intergroup hostility and perpetuation of prejudice. Specifically, in line with prior research that demonstrated the reciprocal relationship between meta-dehumanisation and dehumanisation (Kteily et al., 2016), Pavetich and Stathi hypothesised and found an inverse relationship between meta-humanisation and reciprocal outgroup humanisation (see also Kteily et al., 2016, Study 6). Evidence from both non-Muslims (religious majority) and Muslims (religious minority) in the United Kingdom and Canada showed that priming meta-humanisation increases reciprocal humanisation of the outgroup, which in turn reduces prejudice. Of note, these effects held when controlling for intergroup contact. This indirect effect of meta-
humanisation on prejudice was moderated by intergroup threat, which was measured (Studies 2a, 2b, and 2c) and manipulated (Studies 3a and 3b). The results overall indicated that the indirect effect of meta-humanisation on prejudice was significant predominantly *only* under high intergroup threat.

**Figure 3 approximately here**

Overall, perceiving that one’s ingroup has been afforded human qualities by an outgroup reduces outgroup prejudice via increasing the humanity attributed to the outgroup, even when intergroup threat is high. Preliminary evidence on meta-perceptions, and specifically on the meta- (de)humanisation continuum, demonstrates that meta-humanisation increases the attribution of humanity to the outgroup. More research is essential to allow confidence in the generalisation of the results in distinct contexts and different, (even) more conflicting intergroup relations. There is also a critical need to explore in more depth the effects of meta-humanisation on humanity attribution and prejudice among minority groups. Group status may moderate the effects of meta-humanisation, at least in contexts where status and power differences between minority and majority groups are especially pronounced.

*Social categorisation*

For the sake of conducting a thorough review we also acknowledge research using social categorisation to foster humanity attribution although these studies are not part of our own research programmes. We identify two main social categorisation approaches, based respectively on (1) categorisation of ingroup and outgroup under a superordinate category, and (2) multiple categorisation. In line with the basic tenets of the common ingroup identity model (Gaertner & Dovidio, 2000), making a common superordinate category salient allows to perceive ingroup and outgroup individuals as members of the same group, and therefore to assign similar levels of humanity to the groups (Leyens et al., 2007). In their work, McDonald et al. (2015) focused on emotional similarity between groups, which may act as an antecedent of common ingroup identity (Gaertner & Dovidio, 2000). In two experimental studies conducted in Israel, they informed Jewish
Israeli participants that their emotional reaction to an anger-eliciting story was similar to that shown by Palestinian citizens of Israel (Study 1) or Palestinians of the West Bank (Study 2). Results revealed that emotional similarity, compared with control conditions eliciting low emotional similarity, led to greater humanity attribution assessed with a blatant measure. Initial correlational evidence for the effectiveness of common ingroup identity was provided by Gaunt (2009, Study 2) using Arab high-school students in Israel as the participants with Jewish people as the outgroup. Participants who identified more with the superordinate group displayed lower bias in humanity attribution (calculated as the difference in uniquely human emotions attributed to outgroup vs. ingroup members). Other studies found an association between perceptions of common ingroup identity and increased humanity attribution, assessed with measures of uniquely human traits (Capozza, Trifiletti, et al., 2013) or uniquely human emotions (Andrighetto et al., 2012). Using a different operationalisation of common ingroup identity, Miranda et al. (2014) found, in three correlational studies, that Gypsies and immigrants who assimilated more with the majority culture (and perceived to share a common categorisation) reported higher humanity attribution. In contrast with these findings, however, Rohmann et al. (2009, Study 2), who primed French participants with common ingroup (Europeans) versus national ingroup identity, did not find evidence that common ingroup identity reduces infrahumanisation.

The second approach is based on multiple categorisation, a strategy predicated on the notion that individuals can process simultaneously several distinct social categorisations (Crisp & Meleady, 2012). The underlying rationale is that dichotomous ingroup-outgroup distinctions become less meaningful for intergroup judgments with the increase in the number of social categorisations used to evaluate others (Crisp & Hewstone, 2007; Prati et al., 2021). Prati et al. (2016) investigated the multiple categorisation approach in three experimental studies where they varied the intergroup context considered and adopted different measures of humanity attribution (by employing uniquely human emotions or traits). Results revealed that the effects of multiple categorisation on increased outgroup humanity were sequentially mediated by increased
individuation of outgroup members and reduced intergroup threat. Specifically, multiple categorisation leads to a more individuated mode of thinking, shifting attention away from social boundaries and making categorisation a less useful tool to evaluate the social reality. Combining the two social categorisation approaches, Albarello and colleagues conducted a series of studies investigating the additive effect of multiple categorisation and common ingroup identity on humanity attribution. In one study (Albarello & Rubini, 2012), Italian university students were assigned to a single or multiple categorisation condition, where a target was described along one (White/Black) or two categorical dimensions (White/Black, and immigrant). Afterwards, some of the participants were exposed to a human identity prime, by asking them to complete a measure of identification with all humanity, while the remaining participants were not exposed to any prime. Results revealed that although the multiple (vs. single) categorisation condition fostered greater attribution of uniquely human emotions to the target, the most effective condition combined multiple categorisation with common ingroup identity. Similar findings were obtained with a blatant measure of humanity attribution, asking participants to rate whether human rights (taken from the Universal Declaration of Human Rights) applied to the target. These findings were replicated by Albarello et al. (2018).

In sum, there is preliminary evidence showing that social categorisation processes can be tweaked to foster human attribution. However, there are few tests of mediators of the social categorisation effects on humanity attribution. Research in this area would also benefit from the investigation in field settings and across cultures, especially when examining multiple categorisation.

**Outgroup-Independent Strategies**

*Human-Animal Divide*

Other strategies for humanising outgroup members take a more indirect path, not necessarily targeting representations of or feelings toward the outgroup in question. For instance, the Interspecies Model of Prejudice (IMP; Costello & Hodson, 2014a; Hodson, MacInnis, & Costello,
seeks to reduce prejudices toward human outgroups (e.g., immigrants) by undermining or reducing their dehumanisation. It proposes doing so by changing the manner in which people think about human-animal relations. The basic rationale is that animalistic dehumanisation is afforded social value because animals themselves are undervalued relative to humans. To characterise an outgroup as animal-like only has the power to hurt or delegitimise the outgroup to the extent animals are considered ‘inferior’ to humans. The IMP model therefore proposes that when people think about animals as different from and inferior to humans, this facilitates or encourages thinking about one’s human outgroups as animal-like, which in turn promotes prejudice. But this chain of psychological events offers considerable promise relevant to humanising outgroup members if interventions are able to induce thinking in people that animals are similar to humans, which in theory should reduce outgroup dehumanisation (i.e., humanise the outgroup).

Costello and Hodson (2010, Study 1) first examined the basic relations between the key variables in the model to determine whether there is viability in the notion that human-animal relations could impact humanity attribution. This first study involved 70 Canadian university students, who completed measures tapping the extent to which human and animals are similar (vs. different), humanity attribution (of both the uniquely human emotion and trait types discussed previously), and attitudes toward immigrants. We also assessed several potentially relevant individual difference measures, namely social dominance orientation (SDO; the belief that [human] groups should be hierarchically arranged; Sidanis & Pratto, 1999) and universal orientation (an orientation that emphasises the similarities and not differences between people). As expected, we found that those who consider humans and animals as more similar than different not only scored lower in SDO ($r = -.43$) and higher in universal orientation ($r = .38$) but also conceptualised immigrants as significantly more human ($rs = .25$ to .45) and expressed less prejudice toward immigrants ($r = -.43$). In terms of the mediation model, there was support for the IMP model: human-animal similarity predicted lower prejudice indirectly via greater outgroup humanisation (standardised indirect effect = -.10).
Following up on this preliminary support, Costello and Hodson (2010, Study 2) introduced an experimental manipulation, randomly assigning 120 Canadian undergraduates to read one of four scientific-themed articles that either stressed the similarity or differences between humans and animals. The articles also varied in the focal point, that is, whether animals were compared to humans or humans were compared to animals. In keeping with the IMP conceptualisation, it was predicted that stressing animals-as-similar-to-humans (i.e., “elevating” animals toward humans) would result in greater humanisation of a dehumanised human outgroup (immigrants), along with lower prejudice toward the group. For instance, immigrants were considered to have more uniquely human emotions in the animals-are-human-like condition ($M = 5.96$) than when humans-are-animal-like ($M = 5.23$) were primed or in either condition where differences between humans and animals were emphasised ($Ms = 5.04 – 5.05$). Thus, by shaping how people think about animals, and in particular psychologically raising them ‘up’ to humans (boosting their social value), the derogatory and delegitimising power of seeing human outgroups as animal-like was eliminated.

We reasoned that these processes probably develop early given the deeply entrenched undervalueing of animals in most cultures. In several studies Costello and Hodson (2014a) therefore examined outgroup humanisation in 6-10 year old White children in Canada, as a function of human-animal similarity. Prior to this point very little empirical evidence even sought to examine humanity attribution at this age. Both studies examined humanisation in terms of uniquely human emotions and traits, and both involved attitude and humanity attribution measures modified for use by children. This involved using a board with images of people in one location, images of various animals in another, allowing the children to move animals and humans close or far apart when asked about their similarity or difference. The pilot effort (Study 1, $n = 20$) confirmed the viability of the measures, showing that children this age meaningfully dehumanised Black child targets. Of interest to our current discussion, greater human-animal similarity was associated with greater humanisation of Black children with regard to the uniquely human emotions ($r = .47$) and trait-based ($r = .42$) measures. Encouragingly, after watching a 15 minute “Share the World” video on
human-animal similarities for children, human-animal similarity perceptions were substantially boosted ($Ms = 20.20$ vs. $10.75$, $d = .76$). Study 2 ($n = 53$) likewise found that human-animal similarity perceptions predicted greater humanisation of Black children ($r = .45$). Moreover, their parents also showed relations between human-animal similarity and their dehumanisation of Black people ($r = .43$). In a between-subjects design, children who watched the Share the World video (vs. a child’s recycling video) reported significantly less human-animal divide ($Ms = -0.32$ vs. $.31$, $d = .77$). Unfortunately the video was not able to significantly alter humanity attribution to Black children.

Despite the evidence confirming that human-animal similarity is systematically linked to humanity attribution, a follow up study by Costello and Hodson (2014b) showed that people resist and reject this notion. In a sample of largely White university students in Canada ($n = 139$), participants were asked the degree to which a variety of factors caused and solved both dehumanisation of human outgroups and ethnic prejudice. On scales ranging from 1-7, participants overwhelmingly blamed factors such as closed-mindedness ($M = 6.60$ dehumanisation; $M = 6.09$ ethnic prejudice) and believed in solutions such as positive contact with the outgroup ($M = 6.21$ dehumanisation; $M = 6.23$ ethnic prejudice). Strikingly, the only potential cause that was rated significantly below the scale midpoint was the human-animal divide ($M = 2.42$), and the only solutions below the scale midpoint pertained to highlighting animal-to-human similarity ($M = 2.31$) or human-to-animal similarity ($M = 2.14$). Participants failed to see positive links between these variables; moreover, from the full list of potential causes shown to participants, these were the only to be significantly rejected or disavowed. Intriguingly, however, their own human-animal similarity perceptions were positively correlated with humanity attribution ($r = .36$) in this sample. While denying the positive association between these variables, this sample nonetheless showed evidence of a reliable link themselves. Collectively, these studies (Costello & Hodson, 2010, Studies 1-2; Costello & Hodson, 2014a, Studies 1-2; Costello & Hodson, 2014b) reveal consistent evidence that human-animal similarity is naturally associated with greater humanity attribution (in university
students, children, and community adults), and evidence that these human-animal perceptions are malleable and thus strong candidates for interventions. Future research is needed, however, given that such manipulations were able to shape humanity attribution perceptions among university students (Costello & Hodson, 2010, Study 2) but not children (Costello & Hodson, 2014a, Study 2). Future research could also take advantage of the observed links between human ethnic prejudices and speciesism, in keeping with the Social Dominance Human-Animal Relations Model (SD-HARM; Dhont, Hodson, & Leite, 2016). That is, because social dominance orientation (i.e., endorsement of group dominance and hierarchy) underpins the links between ethnic prejudice and speciesism, interventions that focus on reducing social dominance orientation could reduce both forms of bias, and, in theory, could reduce dehumanisation of human groups as a result of reducing speciesism.

Attachment Orientations

According to attachment theory, individuals have an innate attachment system that, by fostering proximity, allows us to receive protection and support from caregivers and close others (Bowlby, 1982). In other words, the attachment system allows people to meet fundamental human needs such as safety, protection, and social affiliation, which from early developmental stages influence how adults experience relations with others. Three main types of attachment have been identified, deriving from child-caregiver relations. The first is secure attachment, when the caregiver is constantly available and supportive, allowing the development of a sense of security and positive relationships with others. When the caregiver is not fully available or supportive, children may develop an anxious attachment style, with the child keeping demanding support. Avoidant attachment occurs when the child is apparently indifferent to the caregiver’s returns after being unavailable (Ainsworth et al., 1978).

Secure, anxious, and avoidant attachment styles, working as internalised working models (Baldwin, 1992), can also be identified in adulthood, with secure individuals maintaining positive representations of self and others, anxious individuals being uncertain regarding the value of the
self and the possibility to develop positive relations with others, and avoidant individuals relying mostly on the self and avoiding close relations with others (see Mikulincer & Shaver, 2016). In other words, secure attachment is regarded as the most adaptive, producing the better outcomes in terms of the development of positive emotional reactions, such as more empathy (Joireman et al., 2002) and trust (Simmons et al., 2009), and better management of threats (e.g., Dewitte et al., 2007). In addition, secure attachment can have beneficial effects in response to stressful events, and in the development of positive relationships with others (Mikulincer & Shaver, 2016; but see Ein-Dor et al., 2010). Note that despite attachment styles being developed during the infancy, they can change in adulthood as a function of relational experiences at older age (Mikulincer & Shaver, 2003; see also Mikulincer & Shaver, 2016).

**Secure attachment and intergroup relations.** Although attachment styles are typically associated with research in the interpersonal domain, they can have an important role when considering relations with unknown others (Belsky, 1999), and thus have relevance to intergroup relations. For instance, secure individuals display lower prejudice compared to insecure individuals (Di Pentima & Toni, 2009). Indirect evidence is provided by Putra, Campbell-Obaid, and Suwartono (2020). They primed individuals with the idea of goodness of humans, consistent with a secure attachment schema where others can be trusted. Results of three experiments revealed that priming led to also view outgroups as being characterised by a good nature and, as a consequence, to include them in a superordinate human group (therefore assigning to them full human status). Supporting specifically the role of attachment styles in contributing to determine the development of positive intergroup relations, Boccato et al. (2015) conducted three correlational studies investigating the relationship between Italians and immigrants from the perspective of Italians. Results revealed positive associations between secure attachment and more frequent and positive contact.

Importantly for our purposes, attachment orientations can be experimentally primed (Mikulincer & Shaver, 2007). Security priming, such as when participants are exposed to stimuli
linked to security or mental representations of caregivers, has beneficial effects on positive emotions such as compassion (Mikulincer et al., 2005) and empathy (Mikulincer et al., 2003). The positive effects of security activation were also found in intergroup contexts, where security primes reduced prejudice and fostered willingness to meet outgroup members (Mikulincer & Shaver, 2001), lowered discriminatory behaviour and improved outgroup attitudes (Boag & Carnelley, 2012). Regarding mediational processes, positive relational outcomes of attachment security overlap with constructs typically associated with more positive intergroup relations. Indeed, the effects of security priming on intergroup outcomes are mediated by higher empathy (Boag & Carnelley, 2016), as well as lower intergroup anxiety and threat (Mikulincer & Shaver, 2001; see also Saleem et al., 2015). Importantly, the effects of security priming were not moderated by individuals' dispositional attachment orientations where examined (Mikulincer & Shaver, 2001).

Indirect evidence for the role of secure attachment in fostering humanity attribution has been provided by Zhang et al. (2015), who tested whether security priming (compared with control conditions) impacts humanity attribution toward others. In three experimental studies with university student participants the authors focused on interpersonal security, implying the sense of being cared for and loved, a construct conceptually similar to secure attachment. The main rationale was that, in contrast to attachment security which typically refers to attachment with close others, interpersonal security would also refer to distant others. Priming security toward others was expected to increase feelings of being connected with all humanity, favouring humanity attribution toward unknown individuals. Results revealed that security priming fostered the attribution of human nature traits (Haslam, 2006) to a woman who had committed a crime (Study 1) and to inhabitants of a fictional city who had been asked to move to another city to supply the labour force (Studies 2 and 3).

**Attachment styles and humanity attribution.** Taking into consideration the literature reported above, we conducted a series of studies with two aims: (1) investigating whether attachment styles are associated with humanity attribution, by considering both dispositional
attachment and the activation of secure attachment; (2) uncover relevant mediating variables. In the first, correlational study (Capozza et al., 2018), we focused on attachment styles as dispositional variables, hypothesising that dispositional security would be associated with greater humanity attribution. We also investigated affective factors, and specifically intergroup anxiety, empathy, and trust, as potential mediators. Indeed, security has been found to reduce feelings of anxiety and threat evoked by the outgroup (Mikulincer & Shaver, 2001). Secure attachment has also been found to be associated more generally to increased compassion toward others, therefore supplying the foundation for testing it as an antecedent of affect (Mikulincer et al., 2005). In turn, the three emotions have been found to be associated with humanity attribution (Capozza, Falvo, et al., 2013).

As the target outgroup, we focused on individuals with intellectual disability, a stigmatised group often denied a fully human status (Capozza et al., 2016; Falvo et al., 2014). Participants were 92 university students without an intellectual disability. The Attachment Style Questionnaire (ASQ) was used to assess adult anxious, avoidant, and secure attachment orientations (Feeney et al., 1994). Humanity attribution was assessed using four uniquely (e.g., rationality) and four non-uniquely (e.g., instinct) human traits, asking individuals to rate outgroup members on each trait. Participants were also administered measures of intergroup anxiety, intergroup trust, and empathy (a scale including both cognitive and affective dimensions), tested as mediators. Regression analyses revealed that secure attachment was positively associated with the attribution of uniquely human traits to the outgroup ($b = .38$); unexpectedly, a positive association for uniquely human traits also emerged for anxious attachment ($b = .43$). Further regression analyses showed that secure attachment was negatively associated with intergroup anxiety ($b = -.34$), while a positive association emerged between anxious attachment and anxiety ($b = .26$). Contrary to predictions, however, none of the indirect effects was significant.

In two experimental studies we further examined the effects of priming secure attachment (Capozza et al., 2021). In order to provide generalisability to the effects of secure attachment priming we focused on two stigmatised and often dehumanised groups: homeless (Study 1) and
Roma people (Study 2). In the first study, among university students \((n = 75)\) secure attachment was primed using stimuli indicating the availability of typical attachment figures, such as parents with a child in their arms, two people in love, grandparents with their grandchildren. We also included two control conditions. In the first, participants were exposed to stimuli depicting single individuals (men and women of different ages, mostly smiling or with a neutral expression), to rule out alternative explanations based on exposure to human beings. We also included a further neutral condition using landscapes. Participants were presented with six uniquely human (three positive and three negative) and six non-uniquely (three positive and three negative) human emotions; they were asked to select those best characterising the outgroup. Results revealed that, although no significant difference emerged between conditions neither for non-uniquely nor for uniquely human emotions, homeless people were attributed more uniquely \((M = 2.68)\) than non-uniquely \((M = 2.16)\) human emotions in the secure attachment prime condition \((\eta^2_p = .07)\); no difference between the two types of emotions emerged in the latter two conditions. In other words, a process of outgroup humanisation was only evident in the secure attachment prime condition.

The second study (Capozza et al., 2021, Study 2) was intended to further show the role of security attachment priming among Italian adults \((n = 242)\) and their reactions to Roma people. To extend and enhance the generalisability of our findings, in addition to varying the target outgroup, we implemented several design features. The study included an experimental (security priming) and a control condition. In the experimental condition, we focused on interpersonal security to prime security (see Zhang et al., 2015): participants were invited to relive a recent episode or interaction with other people that granted them feelings of interpersonal security and warmth. In the control condition, they were invited to remember an episode in which some people turned to them for street directions. The measure of humanity attribution consisted of four uniquely human and four non-uniquely human traits. To further explore humanisation, we included an additional measure of human attribution based on five HN traits (three positive and two negative; see Haslam, 2006). Participants were asked to attribute each trait to the outgroup. Intergroup emotions (intergroup
anxiety, empathy, and trust) were tested as underlying processes. To further explore the effectiveness of security priming, we investigated the moderating role of dispositional attachment orientations (cf. Mikulincer & Shaver, 2001), assessed as in Capozza et al. (2018). Results revealed a main effect of condition: participants assigned all three humanity dimensions to the outgroup more in the security priming ($M = 4.01$) than in the control condition ($M = 3.82$; $\eta_p^2 = .02$). Second, we tested intergroup empathy and trust as mediators of security priming (condition did not affect intergroup anxiety). Results provided evidence for intergroup empathy, but not trust, as mediator of the effect of security priming on uniquely human and HN traits; mediation was nonsignificant for non-uniquely human traits (95% bootstrap CI for intergroup empathy comprised between .001 and .14 for uniquely human traits, and between .01 and .15 for HN traits) (Figure 4). Possibly, feelings of acceptance and support activated by the security priming raised attention to the outgroup, favouring a general increase in trait attribution. It should be noted that intergroup empathy mediated effects on uniquely human, but not on non-uniquely human traits. This finding indicates that affective processes, raised by priming security, foster the attribution of higher humanity; they do not have unspecific effects on all outgroup characteristics. Finally, we did not find moderation by dispositional attachment styles, providing further confidence in the generality of effects by security priming.

Figure 4 approximately here

Research on the effects of secure attachment on humanity attribution is only at its beginning. The studies above generally provide evidence showing that secure attachment can be effective in fostering humanity attribution (Capozza et al., 2018, 2021; Zhang et al., 2015). These findings were observed using both correlational and experimental methodologies, considering both dispositional (Capozza et al., 2018) and situational attachment (Capozza et al., 2021), and different stigmatised outgroups (individuals with intellectual disability, in Capozza et al., 2018; homeless and Roma people, in Capozza et al., 2021), different measures of humanity attribution (attribution of uniquely human traits, in Capozza et al., 2018, and in Capozza et al., 2021, Study 2; attribution of uniquely
human emotions, in Capozza et al., 2021, Study 1). Moreover, empirical findings provided initial evidence for intergroup emotions, and specifically for intergroup empathy (Capozza et al., 2021, Study 2), as mediators of the effects of secure attachment.

**Future Directions**

*Identifying Additional Strategies*

In addition to the five strategies that we identified, there is some evidence that other strategies may be viable as well. Prati et al. (2015) conducted two experiments where participants were asked to engage in counter-stereotyping thinking. Results revealed that this strategy promoted humanity attribution toward a wide range of outgroups. Notably, these were unrelated to outgroups used in the experimental manipulation, with effects being mediated by lower tendency to rely on heuristic thinking. Other research can explore different alternative methods of assessing humanisation. For instance, Hodson and Doucher (2020) employed subtle changes in language to impact the perceived agency and experience (indicators of humanisation) in social targets (see also E. Cooley et al., 2017). Specifically, targets described as *people in a group* (group composition), as opposed to a *group*, were rated as higher in mind perception, captured by greater felt experience ($M_s = 82.68$ vs. $67.42$) and agency ($M_s = 81.95$ vs. $70.37$). This suggests that subtleties in describing social targets can meaningfully humanise those targets, requiring very little effort or cost. Saguy et al. (2015) tested awareness of intergroup help as an effective strategy to humanise the outgroup in conflicting situations. In one experiment, they showed that being aware that their ingroup helps the outgroup led Israeli-Jews to humanise Palestinians; Study 2 found that this effect is specific to ingroup helping, disappearing when intergroup helping is offered by a third party.

*Investigating processes specific to dehumanisation*

One important question that remains unanswered in the present review is whether there are underlying processes which differentiate prejudice reduction from dehumanisation reduction. Existing studies do not allow us to identify such processes. Future research should include prejudice measures along with dehumanisation, and test the relative strength of mediators for prejudice or
dehumanisation respectively. Our first consideration relates to the existence of such processes. There is evidence that contact effects on prejudice or dehumanisation are mediated by group representations and identification with common identity (Gaertner & Dovidio, 2000). It was also found that contact indirectly reduces prejudice via an increase in identification with humanity (Sparkman & Hamer, 2020). Although this process can also lead to prejudice reduction, we believe that the conceptual correspondence between outgroup humanisation and inclusion of the (humanised) outgroup in a common human identity makes this process primarily relevant for dehumanisation. The second consideration concerns when it is more likely to find specific underlying processes, that is, which processes relate to which forms of dehumanisation. We believe that, rather than merely being specific to dehumanisation, some mediators may be particularly relevant to the reduction of specific forms of dehumanisation. For instance, intergroup anxiety can be more relevant to the reduction of blatant rather than subtle forms of dehumanisation, given that a blatantly dehumanised outgroup can be perceived as highly threatening. Also, mediators can be specific to animalistic versus mechanistic dehumanisation. For instance, cognitive processes may be especially relevant to animalistic, or to mechanistic dehumanisation. It is also possible that processes underlying reduction of prejudice and dehumanisation overlap: given the absence of investigation of underlying factors specific to dehumanisation, we believe this is an important field for future research.

**Extending our Understanding of the Five Different Strategies Highlighted**

With respect to intergroup contact, research on several aspects is scarce or missing. First, research has largely overlooked the examination of participants from low-status groups. This is a very relevant point, since lack of perceived outgroup humanity from the low-status group may prevent social integration within the larger society. Second, research has mainly focused on positive or neutral contact, while negative contact has received less attention: given its role in shaping outgroup attitudes, also in interaction with positive contact (Schäfer et al., 2021), it is important to test negative contact when considering dehumanisation. Third, many studies are correlational in
nature, highlighting the need for experimental and longitudinal evidence. Fourth, experimental research (especially experimental and longitudinal) on the effects of contact on humanity attribution in naturalistic contexts is basically non-existent, leaving doubts about the effectiveness of structured contact interventions. Fifth, research has concentrated on the examination of subtle forms of dehumanisation; less is known about the effects of contact on blatant dehumanisation, especially when considering underlying processes. More generally, research on the mediators of humanity attribution is scarce. Finally, research failed to investigate moderators. As an example, although contact has been shown to be more effective for more prejudiced people (R. N. Turner et al., 2020), overcoming blatant or subtle dehumanisation may be more difficult.

With respect to meta-humanisation, there is the need for a more thorough examination of the processes that explain the meta-humanisation-to-humanisation cycle. For instance, the cycle may be instigated by a reciprocity mechanism, as well as by psychological processes specifically related to intergroup relations. As an example, fear of being rejected is one of the barriers between groups (Stathi et al., 2020): being granted full human status may increase the expectation of being accepted (or reduce perceived discrimination, another obstacle to positive intergroup relations; Bagci, Celebi, & Karakose, 2017), with this more positive attitude predicting greater humanity attribution. A further psychological process candidate is morality. Research shows that morality has an important role in impression formation, and that perceptions of outgroup morality shape outgroup attitudes (Brambilla, Sacchi, Pagliaro, & Ellemers, 2013). Meta-humanisation may therefore foster perceptions of similarity and that the outgroup has moral qualities to the same extent as the ingroup. Such acquired morality may be the proxy for attributing full human status to the outgroup. Importantly, attention should be paid to differentiating morality as a mediator from the morality component that characterises some measures of humanity attribution. A further area of development concerns the conditions that enhance or inhibit such meta-humanisation-to-humanisation cycle. In particular, intergroup threat deserves further investigation. Although we found that the cycle can also work in face of high threat (Pavetich & Stathi, 2020), a fuller range of threats warrants
investigation. As an example, threat from the outgroup to the moral qualities of the ingroup may raise a barrier that breaks the cycle. In contrast, evidence that the outgroup possesses moral qualities can instead boost the cycle.

For social categorisation, future research might also benefit from a more thorough examination of mediators. Consistent with our argument for meta-humanisation, outgroup morality attribution may play a relevant role. Perceptions of belonging to a superordinate group can enlarge one’s moral community (Opotow, 1990), allowing the extension of perceived morality reserved to ingroup members to members of the former outgroup, in turn leading to humanity attribution. In this sense, distinct common identities may have conceptually distinct effects. For instance, effects via outgroup morality attribution may be more likely when the common identity is represented by the human group; when different subgroup identities are also salient, such a process may be inhibited. Multiple categorisation may work along different psychological processes. For instance, Prati et al. (2016) showed that multiple categorisation was associated with humanity attribution via greater individuation of outgroup members. If the mechanism driving multiple categorisation implies lower salience of group membership, then its potential for generalisation of effects may be low (unless it can promote a new way of thinking free of stereotypes). Therefore, it may be more effective in highly conflictual contexts characterised by blatant dehumanisation, where the immediate need is to reduce conflict between specific parties.

We believe research on human-animal similarity can develop in several directions. There is a need to conduct field studies, and to further investigate how manipulations of human-animal similarity can foster humanity attribution in children. Researchers can also focus on the processes through which human-animal similarity promote humanity attribution. Once again, we believe increased attribution of morality to the outgroup can play a relevant role. But there may be other paths, also potentially unrelated to intergroup relations. Borrowing the concept of tertiary transfer effect from the contact literature (Hodson, Crisp, Meleady, & Earle, 2018; Meleady, Crisp, Hodson, & Earle, 2019), human-animal similarity, which is an outgroup-independent strategy, can lead
individuals to broaden their horizons, challenge their worldviews, use different lens to interpret social reality. In other words, fostering human-animal similarity can act as an agent of cognitive liberalisation.

Research on security attachment to promote humanity attribution is still at its beginning. One important step is start identifying mediators. Initial empirical evidence is mixed, and only provided some support for intergroup empathy (but not for intergroup anxiety or trust). Possibly, mixed findings on the mediators of attachment orientations are due to the fact that scholars focused on variables related to specific outgroups, instead of looking for more general processes, not necessarily related to intergroup relations (consistent with our classification placing security attachment into outgroup-independent strategies). Future research might also investigate further affective factors, such as threat, and eventually explore the mediating role of cognitive variables. For instance, security activation may provide more confidence in relations with outgroup members (R. N. Turner & Cameron, 2016), contributing to reducing negative intergroup expectations.

Research can also investigate the conditions where security activation is more effective. For instance, because secure attachment helps people facing stressful events (Mikulincer & Shaver, 2007), it may produce stronger effects in conflictual or threatening situations, or on variables implying an especially strong conflict such as blatant dehumanisation.

With respect to research on human-animal similarity and security attachment, it is worth noting that, at least in some conditions, social connection can increase dehumanisation of others (Waitz & Epley, 2012). In addition, lack of secure attachment can increase humanity attribution to non-human others (Waytz et al., 2013). In other words, while secure attachment can increase humanity attribution to human outgroups, and in some cases promote dehumanisation of distant others, lack of attachment can foster humanity attribution to non-human outgroups. These findings reveal that research on attachment orientations and human-animal similarity are interrelated and the effects of attachment orientations on ingroup, outgroups, and non-human outgroups may be complex, requiring further investigation.
**Building a Unitary Approach that Synthesises the Five Strategies**

Although we treated the five strategies as independent, we see considerable value in considering their integration. In theory they can be implemented simultaneously or sequentially in time. Presumably they can also influence each other. For instance, activating security or meta-humanisation (or both) might prepare prejudiced individuals for contact. However, strategy integration may be driven by the specific processes activated by the different strategies. As we have argued, intergroup contact, meta-humanisation and social categorisation are largely outgroup-specific strategies. Their implementation is rooted in a relationship with a specific outgroup, increasing the likelihood that (if effective) they will boost humanity attribution primarily toward that outgroup (although we do not exclude generalisation effects). This has at least two consequences. First, underlying processes will be most relevant to the specific intergroup relationship under investigation. Candidate mediators are the constructs identified in the contact research, such as intergroup emotions or cognitive representations. Second, such strategies may be especially relevant in non-conflictual contexts, where individuals may raise psychological barriers if explicitly faced with strategies focused on the “enemy.”

Human-animal similarity and security activation in relation to attachment orientations can be considered outgroup-independent strategies, since their implementation does not refer to a specific (human) outgroup. In this case, mediating processes can extend to psychological constructs less related to specific outgroups, such as personality variables, variables related to one’s ingroup. For instance, security activation may allow a more confident way to interact with others, promoting increases in individuals’ agreeableness. Human-animal similarity can lead individuals to reconsider their ingroup, consistent with the process of dep provincialisaton (see Pettigrew, 1998). A less provincial view of the ingroup implies that there may be alternative ways to behave and interpret social reality. These processes in turn should favour humanity attribution and prejudice reduction. A second relevant theoretical difference between outgroup-specific and outgroup-independent strategies is that outgroup-independent strategies may be more likely to favour humanity attribution
toward a wide range of outgroups. If individuals embrace the idea that animals are similar to humans, they may potentially apply the greater humanisation to many stigmatised categories who generally suffer (animalistic) dehumanisation. Similarly, promoting a more secure way to interact with others can potentially extend to all individuals’ interpersonal and intergroup relations. Being independent of the intergroup context, outgroup-independent strategies may be effective in enhancing humanity attribution among people more prone toward prejudice who might otherwise resist interventions that appear intergroup in nature (see Hodson et al., 2013).

Based on the above, strategy combination can be based on the different and potentially complementary processes that they activate. Strategies that activate the same underlying processes can be potentially redundant and do not increase their general effectiveness. For instance, if two strategies (e.g., intergroup contact and common ingroup identity) foster intergroup empathy, then using both may not constitute an advantage in terms of effectiveness. In order to combine strategies with this aim, it is important to clarify the potential. We also propose a different rationale through which the combination of strategies can lead to greater effectiveness, drawing on the work of Akrami et al. (2011), who argued that prejudice toward a target group can be differentiated into a specific and a common component. Whereas the specific component captures the attitude toward the specific group, the common component consists in variance of prejudice that the individual has toward all groups. Therefore, for instance, prejudice toward immigrants can be differentiated into a specific component, that may be a function of negative stereotypes characterising immigrants and the threat associated with them, and in a common component (or generalised prejudice). This latter component does not depend on the characteristics of the immigrant group, but represents the variance of prejudice shared with other groups, including for instance immigrants, gay people, ethnic groups, etc. Therefore, the common component is more likely to be impacted by ideological variables such SDO (see Hodson et al., 2017), or variables unrelated to the specific group like personality variables (Akrami et al., 2011).
We argue that a similar distinction can emerge with respect to humanity attribution. It is possible that humanity attribution can be differentiated in a specific and a common component, with the first being mostly a function of features relevant to the specific outgroup, and the second mostly affected by ideological or personality variables. In addition, the common component may be a function of variables referred to the ingroup (e.g., related to the deprovincialisation process proposed by Pettigrew, 1998): changes in how the ingroup is appraised are not necessarily related to a specific intergroup relation, therefore they may have consequences for a wide variety of outgroups (for a similar argument related to contact and generalised prejudice, see Vezzali & Stathi, 2021, Chapter 6; Vezzali et al., 2021). Returning to our distinction, while outgroup-specific strategies can mostly impact on the specific component of humanity attribution (although contact can also lead to generalised prejudice; Vezzali & Stathi, 2021, Chapter 6; Vezzali et al., 2021), outgroup-independent strategies could mainly change the common component of humanity attribution. Combining strategies with this rationale in mind could allow to “attack” humanity attribution on different levels, increasing the likelihood of the success of an intervention.

**Fighting Dehumanisation with Practical Interventions**

Although the presented strategies can be effective in reducing dehumanisation, ecological validity provided by studies in the field is rare. The only two studies that we are aware of, described in the previous sections, concerned (imagined) contact (Vezzali et al., 2012), and human-animal similarity (Costello & Hodson, 2014a), both showing an increase in humanity attribution among young children. However, only the study by Costello and Hodson (2014a) was specifically focused on boosting attribution of humanness to animals, since the video shown as the manipulation to children discussed similarities between humans and animals. In contrast, the imagined contact intervention (Vezzali et al., 2012) asked children to imagine an outgroup person, with outgroup humanity attribution as one of the dependent variables; in other words, this study was not specifically focused on presenting the outgroup person as a human or in human terms.
With respect to contact, practical interventions might focus on facilitating reciprocal disclosure of humanlike attributes, such as uniquely human emotions, during contact. For instance, participants from two different groups might be involved in a common task where they are asked to disclose to each other the emotions experienced while watching a video showing violence against outgroup victims, or an outgroup person needing help. Interventions on meta-humanisation can capitalise on videos that show outgroup members expressing their attitudes and emotions toward the participants’ ingroup, providing this way ‘evidence’ that the outgroup humanises the ingroup. Using social categorisation in interventions, individuals can be presented with pictures of individuals from different groups and of different animals, and asked to engage in multiple categorisation tasks revealing the overlap between human groups, and at the same time highlighting differences from the non-human groups (multiple categorisation). Again, participants can be invited to include humans from different groups into a single human category against a non-human category (e.g., robots), with the task of identifying how humans differ from robots and how they can help each other because they are motivated (rather than ‘programmed’) to do it. To implement human-animal similarity interventions, participants can be invited to write down what unites humans and animals. Finally, priming secure attachment to individuals, followed by activities that show different human groups providing social connection and reciprocal help, can be used to implement secure attachment strategies in the field.

Importantly, these strategies should not be considered separately, but as a set of tools that can be used one at a time or altogether, depending on the specific situation. For instance, the human-animal similarity strategy may be less useful in reducing mechanistic dehumanisation; outgroup-independent strategies may encounter less resistance in conflictual settings. Integrating strategies can be the most effective way to foster humanity attribution. For example, given their outgroup-independent focus, human-animal similarity and security attachment can be used as preparatory strategies, whereas contact (eventually paired with social categorisation as members of
a common group) can facilitate reduction of dehumanisation toward a specific target group.

**Conclusion**

We presented evidence of five different strategies that have been found to be effective in fostering humanity attribution. Note that research on the strategies that reduce dehumanisation is relatively nascent relative to those studying its exacerbating factors. One consequence is the general lack of information on the processes underlying changes in human attribution, and future research should shed light on these mediators in order to plan effective interventions that can be adapted to the different intergroup contexts. It is worth noting that our conclusions require caution because of some important limitations of the presented evidence: evidence is mainly correlational (especially for the contact strategy) and sometimes obtained with small samples; and evidence for the other strategies (especially for attachment orientations) is still at its infancy. In our review we have differentiated these strategies into outgroup-specific and outgroup-independent strategies. The main advantages of such a distinction pertain to the variety of outgroups potentially impacted by the strategies and the possibility of better differentiating underlying processes. These two aspects may be important to combine in order to maximise their efficacy. In addition to the theoretical relevance of these conclusions, there are practical consequences. Given widespread dehumanisation in contemporary society, both at the subtle and blatant level, interventions should be focused on targeting humanity attribution specifically. Multicomponent interventions may be best equipped to impact the different facets of dehumanisation: the present review offers a panoramic view on the most effective strategies offered by research so far, and provides indications on how to combine them in order to construct theoretically-driven interventions in the field that can inform social policies.

**Footnotes**

1. In the remaining of this article we will use the term ‘humanity attribution’ or ‘dehumanisation’ to refer to humanity attribution toward or dehumanisation of the *outgroup*.

**References**


Brambilla, M., Sacchi, S., Pagliaro, S., & Ellemers, N. (2013). Morality and intergroup relations:

Threats to safety and group image predict the desire to interact with outgroup and ingroup members.

http://dx.doi.org/10.1016/j.jesp.2013.04.005


http://dx.doi.org/10.1002/ejsp.384


http://dx.doi.org/10.1177/0146167220949004


http://dx.doi.org/10.1371/journal.pone.0170554


http://dx.doi.org/10.4473/TPM20.4.6


http://dx.doi.org/10.1080/00207594.2012.688132


http://dx.doi.org/10.1111/bjso.12016


Ein-Dor, T., Mikulincer, M., Doron, G., & Shaver, P. R. (2010). The attachment paradox: How can so many of us (the insecure ones) have no adaptive advantages? *Perspectives on Psychological Science, 5*, 123-141. http://dx.doi.org/10.1177/1745691610362349


http://dx.doi.org/10.1080/10463280801981645


http://dx.doi.org/10.1111/j.1467-9280.2007.01962.x


http://dx.doi.org/10.1177/1745691617752324


http://dx.doi.org/10.1177/1368430220979035


http://dx.doi.org/10.1111/j.1751-2409.2007.00006.x


http://dx.doi.org/10.1177/1368430215595107


Miranda, M., Gouveia-Pereira, M., & Vaes, J. (2014). When in Rome...Identification and acculturation


http://dx.doi.org/10.1080/10463280601095240


http://dx.doi.org/10.1002/job.585


https://doi.org/10.1016/j.ijintrel.2020.08.007


Vorauer, J. D. (2013). Getting past the self: Removing evaluative concerns as an obstacle to positive intergroup contact effects. In G. Hodson & M. Hewstone (Eds.), Advances in intergroup contact (pp. 23-48). Psychology Press.


Figures captions

Figure 1. Examples of manipulations used in Capozza et al. (2017, Study 2). In the experimental and neutral stimuli control condition, using the keypad arrows (according to the manikin position), participants were asked to move the manikin (representing the self) upwards (row A) or downwards (row B) to approach the target stimuli, presented in the middle of the computer screen. The target stimuli were Moroccan (outgroup) faces in the experimental condition and images of furniture in the control condition. Geometrical figures (ovals) were also used: participants were asked to move the manikin away from them. In the sideways-control condition, participants were asked to move the manikin toward the right in response to the target stimuli (Moroccan faces), presented in the middle of the screen (rows A and B). Participants also had to move the manikin toward the left when the two ovals were shown.

Figure 2. Mediation model investigating the indirect effect of imagined contact on attribution of uniquely human emotions to outgroup members via outgroup trust. Only significant paths are reported (standardised coefficients). Source: Vezzali et al., 2012. **p < .01. ***p < .001.

Figure 3. Conditional model showing the moderating role of threat on the indirect effect of meta-humanisation to prejudice via outgroup humanisation (Study 2c – Muslim sample). Only significant paths are reported (unstandardised coefficients). Source: Pavetich & Stathi, 2021b. ***p < .001.

Figure 4. Mediation model investigating the indirect effect of security priming on attribution of uniquely human traits and human nature traits via intergroup emotions. Only significant paths are reported (unstandardised coefficients). Findings derive from distinct linear regressions, one for each dependent variable, and are presented in a single figure with descriptive purposes. Source: Capozza et al., 2021, Study 2. *p < .05. ***p < .001.
Experimental condition and neutral stimuli control condition

Manikin presented under the stimulus

A

Manikin presented above the stimulus

B

Sideways-control condition

Manikin presented under the stimulus

A

Manikin presented above the stimulus

B
Figure 2

Imagined contact (vs. control) → Outgroup trust → Uniquely human emotions

0.65**

0.60**

-0.20
Figure 3

Threat -> Outgroup humanisation: -12.05***

Meta-humanisation -> Outgroup humanisation: 8.39***

Outgroup humanisation -> Prejudice: -0.01***
Security priming (vs. control) → Outgroup empathy → Uniquely human traits

Outgroup trust → .13*

.35*

.15*

.29***

.14*