Cognition 212 (2021) 104682

Contents lists available at ScienceDirect

Cognition

journal homepage: www.elsevier.com/locate/cognit

No convincing evidence outgroups are denied uniquely human characteristics: Distinguishing intergroup preference from trait-based dehumanization

Florence E. Enock *, Jonathan C. Flavell, Steven P. Tipper, Harriet Over

University of York, United Kingdom

ARTICLE INFO

Keywords:
Dehumanization
Intergroup bias
Prejudice
Social cognition

ABSTRACT

According to the dual model, outgroup members can be dehumanized by being thought to possess uniquely and characteristically human traits to a lesser extent than ingroup members. However, previous research on this topic has tended to investigate the attribution of human traits that are socially desirable in nature such as warmth, civility and rationality. As a result, it has not yet been possible to determine whether this form of dehumanization is distinct from intergroup preference and stereotyping. We first establish that participants associate undesirable (e.g., corrupt, jealous) as well as desirable (e.g., open-minded, generous) traits with humans. We then go on to show that participants tend to attribute desirable human traits more strongly to ingroup members but undesirable human traits more strongly to outgroup members. This pattern holds across three different intergroup contexts for which dehumanization effects have previously been reported: political opponents, immigrants and criminals. Taken together, these studies cast doubt on the claim that a trait-based account of representing others as 'less human' holds value in the study of intergroup bias.

1. Introduction

In a social world characterised by the continued prevalence of prejudiced attitudes, systematic discrimination and social division, understanding the psychological causes of intergroup bias is essential. The construct of dehumanization is commonly invoked to explain intergroup biases. Although there are different conceptions of dehumanization within the literature, proponents of the view typically hold that some outgroup members are perceived as less human than are ingroup members (Harris & Fiske, 2006; Haslam, 2006; Leyens et al., 2000, 2001). The claim that dehumanization is associated with intergroup harm has been extremely influential and is broadly accepted in many areas of social psychology, social neuroscience and philosophy (Haslam, 2019; Haslam & Loughnan, 2014; Smith, 2011, 2016; Tirrell, 2012). For example, Haslam recently noted "*The research literature reveals so many associations between dehumanization, assessed in a wide variety of ways, and various forms of maltreatment that strong links between dehumanization and violence are difficult to deny*" (Haslam, 2019, pp. 134–135).

Psychological models tend to propose that dehumanization exists along a continuum (Harris & Fiske, 2006; Haslam, 2006; Leyens et al., 2000, 2001). In cases of extreme dehumanization, outgroup members may be likely to fall victim to severe harm such as genocide, torture and rape (Haslam, 2019; Haslam & Loughnan, 2016). Evidence for this type of dehumanization has typically been drawn from historical documents. For example, propaganda in which the victims of genocide are compared to rats, lice or snakes (Haslam & Loughnan, 2014; Smith, 2011, 2014, 2016; Tirrell, 2012). Social psychologists and social neuroscientists propose that, in more subtle forms of dehumanization, outgroup members are perceived as somewhat less human than are ingroup members (Harris & Fiske, 2006; Haslam, 2006; Leyens et al., 2003; Leyens, Demoulin, Vaes, Gaunt, & Paladino, 2007). Evidence for more subtle forms of dehumanization has typically come from lab-based research. Subtly dehumanized groups are thought to be shown less empathy, less forgiveness for perceived wrongdoing, and are less likely to receive help (Haslam & Loughnan, 2014, 2016; Vaes, Leyens, Paola Paladino, & Pires Miranda, 2012; Vaes, Paladino, Castelli, Leyens, & Giovanazzi, 2003; Vaes, Paladino, & Leyens, 2002). These forms of dehumanization are thought to be widespread in society (reviewed in Haslam & Stratemeyer, 2016; Leyens, 2009; Vaes et al., 2012).

More recently, researchers have attempted to bridge the gap between

* Corresponding author at: Department of Psychology, University of York, York YO10 5DD, United Kingdom.
E-mail addresses: Florence.enock@York.ac.uk (F.E. Enock), Jonathan.Flavell@York.ac.uk (J.C. Flavell), Steven.Tipper@York.ac.uk (S.P. Tipper), Harriet.Over@York.ac.uk (H. Over).

https://doi.org/10.1016/j.cognition.2021.104682
Received 8 October 2020; Received in revised form 22 February 2021; Accepted 15 March 2021
Available online 24 March 2021

the study of extreme dehumanization and its more subtle psychological forms by developing a blatant dehumanization scale (Kteily, Bruneau, Waytz, & Cotterill, 2015). In studies using this method, participants are asked to placedifferent social groups on a scale depicting silhouettes ranging from early human ancestors reminiscent of apes through to modern humans. Empirical research has shown that certain outgroups are rated as less human-like on this scale (Bruneau, Kteily, & Laustsen, 2018; Kteily et al., 2015). For example, in a US sample, Arabs and Muslims were amongst groups rated as significantly less evolved than a group labelled Americans. Furthermore, ratings on this scale were associated with wider attitudes. For example, the extent to which Arabs were 'blatantly dehumanized' predicted reduced support for Arab immigration, lower empathy for an Arab individual, and even endorsement of direct harm towards Arabs (Kteily et al., 2015).

Three psychological models of subtle dehumanization have been particularly influential. A social neuroscience perspective proposes that dehumanization is best characterised as a lack of mental state attribution. According to this perspective, to the extent a group is dehumanized, they are viewed as lacking beliefs, desires and intentions (Harris & Fiske, 2006; Harris & Fiske, 2011). Leyens and colleagues (Leyens et al., 2000, 2001) offer an alternative characterisation, known as infrahumanisation theory, that focuses on particular types of mental state attribution. According to this view, subtle dehumanization occurs when outgroup members are perceived to experience complex secondary (uniquely human) emotions to a lesser extent than do ingroup members (e.g., Demoulin, Pozo, & Leyens, 2009; Leyens et al., 2003; Leyens et al., 2007; Vaes et al., 2012).

A third highly influential account of dehumanization is known as the dual model of dehumanization and focuses on trait attributions (Haslam, 2006). While this model hypothesises that dehumanization occurs in both interpersonal and intergroup processes, our primary interest in this paper is the predictions it makes about intergroup bias. The dual model builds on infrahumanisation theory to offer an account in which there are two different forms of (or routes to) dehumanization. Groups are animalistically or mechanistically dehumanized to the extent they are denied particular character traits.

One advantage of the dual model over other theories of dehumanization is that it is based on empirical research into the lay characterisation of the concept human. To determine traits most associated with the concept of humanness, participants were asked to rate the extent to which eighty pre-defined personality traits were uniquely human ('this characteristic is exclusively or uniquely human: it does not apply to other species') or human nature ('this characteristic is an aspect of human nature') (Haslam, Bain, Douge, Lee, & Bastian, 2005). On the basis of participants' answers, two senses of humanness were proposed: uniquely human characteristics, which separate humans from other species, and human nature characteristics, which are supposedly typically or essentially human. Though human nature traits were initially defined as those that are fundamental to humans and independent of any comparison category, they are also proposed to be akin to those that distinguish humans from robots and other machines (Haslam, 2006). Discussing the model, Haslam notes "*The animalistic form of dehumanization rests on a direct contrast between humans and animals, but in the mechanistic form, although the relevant sense of humanness is non-comparative (HN), humans can be contrasted with machines. The shared, typical, or core properties of humanness are also those that distinguish us from automata*" (Haslam, 2006, p. 258). According to this model, uniquely human characteristics can be summarised as civility, refinement, moral sensibility, rationality, and maturity. A denial of these traits is considered animalistic dehumanization, a perception of another person or social group as uncultured, coarse, amoral, irrational, and childlike. Human nature characteristics can be summarised as emotional responsiveness, interpersonal warmth, cognitive openness, agency, and depth. A denial of these traits is considered mechanistic dehumanization, a perception of another person or social group as inert, cold, rigid, passive, and superficial (Haslam, 2006).

Empirical work examining the dual model of dehumanization suggests that some social groups may be animalistically dehumanized whilst others may be mechanistically dehumanized. For example, Loughnan and Haslam (2007) suggested, based on implicit measures, that participants implicitly associated artists more strongly with human nature traits (such as curious and friendly) than uniquely human traits (such as organised and polite), whilst the reverse was true for businesspeople. Further, nonhuman animal terms (such as cattle and primates) were associated with artists more than businesspeople, while automata terms (such as android and computer) were associated more with businesspeople than artists. Loughnan and Haslam argue this shows subtle animalistic dehumanization of artists and subtle mechanistic dehumanization of businesspeople. Other research has suggested that both Anglo-Australian and Ethnic-Chinese participants in Australia tended to mechanistically dehumanize Ethnic-Chinese people, ascribing them human nature qualities to a lesser extent than Anglo-Australians, but tended to animalistically dehumanize Anglo-Australians, ascribing them uniquely human qualities to a lesser extent than Ethnic Chinese people. The researchers proposed that different cultural groups may subtly dehumanize an outgroup along one dimension of humanness, but subtly dehumanize an ingroup along the other (Bain, Park, Kwok, & Haslam, 2009). In related work, Italian participants seemed to animalistically dehumanize Haitians but mechanistically dehumanize Japanese people. These two respective forms of dehumanization purportedly decreased willingness to help in relation to the earthquakes in both countries (Andrighetto, Baldissarri, Lattanzio, Loughnan, & Volpato, 2014).

Recently, the logic underlying the dual model of dehumanization has been called into question (Bloom, 2017; Lang, 2010, 2020; Manne, 2016, 2018; Over, 2020a; Smith, 2014, 2016). Over (2020a, 2020b) points out that the traits thought to characterise human uniqueness and human nature are typically socially desirable in nature. To be conceived of as civilised, rational, warm and cognitively open, for example, is generally positive. Apparent evidence for the dual model of dehumanization may, then, be better explained in terms of ingroup preference and stereotyping. Importantly, antisocial traits such as jealousy, arrogance and selfishness also appear to be characteristic of humans but seem unlikely to be attributed more strongly to ingroup than outgroup members (Manne, 2016; Over, 2020a; Over, 2020b).

Prior work that has measured outgroup dehumanization based on the dual model has not rigorously separated human specific characteristics from ones that are perceived as socially desirable. For example, one study measured dehumanization of criminals by having participants rate different kinds of offenders (i.e. white collar, violent or child molesters) on statements assessing human nature and human uniqueness. Human nature statements included 'I felt like the person in the story was emotional, like they were responsive and warm', and 'I felt like the person in the story was mechanical and cold, like a robot' (reversed). Uniquely human statements included 'I felt like the person in the story was rational and logical, like they were intelligent' and 'I felt like the person in the story lacked self-restraint, like an animal' (reversed). These measures claimed to show criminals to be dehumanized differentially based on crime, and that dehumanization predicted attitudes such as how severely the offender should be punished (Bastian, Denson, & Haslam, 2013). However, because the human characteristics were all positive, desirable ones, it is difficult to separate 'dehumanization' from dislike.

Related research measured animalistic dehumanization of sex offenders by asking participants to indicate the extent to which human words (humanity, person, people, civilian) and animal words (creature, beast, animal, mongrel) described rapists and paedophiles. Here, the extent to which people associated the offenders with animal words predicted attitudes such as reduced support for rehabilitation and longer punishment sentences (Viki, Fullerton, Raggett, Tait, & Wiltshire, 2012). A similar measure suggested that the extent to which Christians animalistically dehumanized Muslims predicted advocacy of torture

towards Muslim prisoners of war (Viki, Osgood, & Phillips, 2013). These tests have also been used to infer dehumanization of political opponents and gay men (Fasoli et al., 2016; Pacilli, Roccato, Pagliaro, & Russo, 2016). However, as noted before, word valence was not appropriately controlled for. Describing someone as 'beast' or 'mongrel' is usually more derogatory than describing someone as 'person' or 'civilian'.

Mechanistic dehumanization has been measured by asking participants how compassionate they believe outgroup members to be. Amongst samples of Palestinians and Jewish-Israelis, this measure predicted support for peaceful compared to violent conflict resolution strategies (Leidner, Castano, & Ginges, 2013). However, compassion is undoubtedly a prosocial and desirable quality. It is conceivable that people who more strongly view outgroup members as having compassion also like them more.

1.1. The present work

These critiques suggest a plausible alternative account of the existing data that appear to support the dual model: outgroups *are* perceived to possess uniquely human attributes to a similar extent as ingroups, but that these attributes are typically negative and antisocial in character. In this paper, we empirically assess the claims of the dual model of dehumanization against this alternative hypothesis.

We first revisit prior accounts that have offered lay conceptions of 'humanness' (Haslam, 2006; Haslam et al., 2005; Haslam, Bastian, & Bissett, 2004) and show that people tend to associate socially undesirable as well as desirable traits with the concept human. In seven experiments, we then examine whether outgroup members are attributed uniquely human traits to a lesser extent even when those attributes are controlled in terms of valence. Across three different intergroup contexts in which dehumanization effects have previously been reported (political opponents, immigrants and criminals), we measure attributions of human specific traits to ingroups and outgroups. By rigorously controlling for trait valence within our experimental designs, we test whether the dual model of dehumanization explains intergroup biases in cognition over and above effects of ingroup favouritism and outgroup dislike.

1.2. Data collection

All studies took place online and were created and administered using Qualtrics (https://www.qualtrics.com), with participants recruited through Prolific (https://www.prolific.co). Informed consent was obtained at the start of each online session according to approved ethical procedures and participants were compensated at an approximate rate of £7.50 per hour. Power analyses were conducted using MorePower 6.0.4. Link to pre-registration documents and raw data for all studies can be found at: https://osf.io/yp8wc/

2. Pretest: How to measure the lay concept of 'human'

We suggest the dual model of dehumanization over-represents positive aspects of humanity and under-represents negative aspects of humanity because of ways in which the content of the lay category 'human' was measured. First, when participants were asked to rate traits for how uniquely or typically human they were, several potentially important undesirable terms likely to be viewed as typically human were omitted from the stimulus set. For example, traits such as jealous, selfish and corrupt were not included. Second, the way in which participants were asked to rate the traits may have inadvertently biased participants to focus on the more positive aspects of humanity. Participants rated 'The extent to which each characteristic was exclusively or uniquely human (does not apply to other species)' and 'The extent to which each characteristic is an aspect of human nature' (Haslam et al., 2005). However, research in the cognitive psychology of categorisation has long demonstrated that attributes that appear typical of a category vary depending on the context such that that our representations are not static, but are constantly changing (Medin & Smith, 1984; Smith & Medin, 1981; Yee & Thompson-Schill, 2016). The set of attributes that are judged unique to humans is likely to depend on the particular comparison. In other words, if we ask participants to compare humans to other species then a different set of uniquely human attributes may emerge than when we ask participants to compare humans to, for example, angels. Furthermore, it is likely that more socially undesirable traits such as jealousy and spite will be listed as uniquely human in the latter context (Over, 2020a).

In order to build understanding of the range of traits that people associate with the category human, we utilised a similar method to prior related work (Haslam et al., 2005) and introduced a broader range of undesirable trait terms. We then asked participants to rate sixty traits on how strongly they applied to humans in three contexts. One scale asked how uniquely human the trait was compared to other species, directly following Haslam et al. (2005). Another asked how uniquely human the trait was compared to robots, following the dual model's suggestion that human nature traits distinguish us from robots and other machines. Haslam and colleagues detail this parallel in much of their prior work, for example, they note "*Our work… proposes two forms of dehumanization, in which people are denied uniquely human attributes and likened to animals, or denied human nature attributes and likened to robots*" (Haslam, Kashima, Loughnan, Shi, & Suitner, 2008, p. 248). A third asked how uniquely human the trait was compared to angels, in order to obtain the more negative or socially undesirable qualities associated with humanness in certain contexts.

Our main aim in this pretest was not to provide an exhaustive account of the lay concept of 'human' as Haslam and colleagues sought to do – a challenging task by any standard. Rather, our aim was considerably more modest – to show accounts of humanness are likely subject to variation depending on context, and to obtain a list of traits rated as highly characteristic of humans as a basis from which to choose stimuli for subsequent studies.

2.1. Methods

2.1.1. Participants

Thirty participants were included (12 female, 18 male), aged between 18 and 51 (Mean age = 27.3, $SD = 7.85$). One participant failed one or more attention checks and their data was excluded and replaced.

2.1.2. Materials

We chose sixty trait words from several sources to reflect a broad range of attributes. An approximate equal number of desirable and undesirable qualities were included. For consistency with previous work, we included terms from Haslam et al. (2004, 2005) and also the five summary 'uniquely human' and 'human nature' qualities along with their opposites from the dual model (Haslam, 2006).

The words we included were: aggressive, arrogant, bitter, calm, capable, civilised, cold, controlling, corrupt, creative, cruel, cultured, cunning, curious, cynical, deceptive, disciplined, dominant, efficient, emotional, energetic, error-prone, ethical, forgiving, generous, gentle, genuine, helpful, honest, humble, immature, impulsive, inflexible, innocent, intellectual, irrational, jealous, kind, knowledgeable, mature, modest, moral, open-minded, passive, rational, refined, selfish, sneaky, sophisticated, spiteful, stingy, stupid, submissive, superficial, trusting, uncultured, unrefined, unsophisticated, warm, wise.

Participants rated these sixty words on three separate sliding scales which asked:

i. How much does this apply to humans compared to other species?
ii. How much does this apply to humans compared to robots?
iii. How much does this apply to humans compared to angels?

Each scale ranged from −50 *Just [comparison category]* to +50 *Just*

humans, with 0 indicating *Equally to [comparison category] and humans*, though participants could not see the scoring numbers. In line with the Dual Model, we were most interested in finding traits perceived to be shared with the comparison category (scoring close to 0) and emotions that were strongly associated with just humans (scoring close to +50). The scales were presented in three separate blocks and the order of completion was counterbalanced. The sixty items within each block were randomised and one attention check per block was also included approximately halfway through, for example *Please indicate 'just humans'*. Additionally, participants completed two short scales to assess attitudes to each comparison category. The order of items participants responded to (other species/angels/robots) was randomised.

Table 1
Pretest results: Means for the trait words on all three scales, ordered from most to least human on each.

Humans v. other species			Humans v. robots			Humans v. angels		
Trait word	*M*	*SE*	*Trait word*	*M*	*SE*	*Trait word*	*M*	*SE*
Cynical	40.9	2.65	Emotional	47.1	1.30	Selfish	43.6	2.11
Corrupt	39.0	4.00	Jealous	40.6	4.32	Corrupt	42.3	2.44
Cultured	38.3	2.91	Selfish	39.1	4.10	Aggressive	41.3	2.51
Open-minded	33.3	4.00	Stingy	38.6	3.39	Jealous	39.8	2.50
Controlling	33.2	3.55	Generous	38.4	3.01	Stingy	39.3	2.75
Arrogant	32.4	3.53	Impulsive	38.3	3.22	Cruel	39.2	2.86
Civilised	31.3	5.07	Cynical	38.1	3.03	Error-prone	38.5	2.96
Sophisticated	30.3	4.43	Open-minded	37.6	3.76	Spiteful	37.8	3.02
Intellectual	29.8	4.24	Warm	37.6	3.01	Bitter	37.0	3.46
Superficial	28.7	5.73	Bitter	37.2	3.56	Cynical	37.0	3.28
Knowledgeable	28.2	4.21	Spiteful	37.0	4.15	Arrogant	35.3	4.43
Bitter	28.1	4.10	Arrogant	34.9	4.07	Immature	34.9	4.04
Stingy	27.2	3.69	Kind	34.3	4.58	Impulsive	34.2	3.71
Moral	27.1	5.21	Cultured	34.2	4.37	Deceptive	33.8	4.41
Refined	27.0	3.90	Sneaky	33.3	4.04	Stupid	33.3	3.05
Selfish	26.4	4.07	Moral	33.1	3.81	Unsophisticated	31.9	3.64
Spiteful	25.7	4.12	Cunning	33.1	3.45	Cunning	30.7	4.01
Humble	25.5	4.13	Cruel	32.8	3.79	Irrational	30.7	4.94
Wise	23.6	4.92	Creative	32.7	4.43	Cold	30.7	3.99
Creative	23.5	4.10	Mature	32.3	3.97	Unrefined	30.4	3.63
Jealous	23.4	3.70	Aggressive	32.2	4.23	Superficial	29.6	4.88
Rational	23.3	5.36	Curious	32.2	4.75	Emotional	28.1	4.48
Emotional	22.2	3.87	Dominant	32.1	3.60	Curious	26.4	4.54
Ethical	21.2	5.78	Forgiving	32.0	4.63	Sneaky	26.3	5.43
Cruel	18.3	4.30	Civilised	31.8	3.94	Energetic	22.6	4.73
Cunning	18.1	4.12	Humble	30.6	4.94	Controlling	20.7	5.25
Mature	17.8	3.94	Deceptive	29.8	3.88	Creative	19.7	5.36
Modest	17.5	5.03	Corrupt	28.8	4.99	Open-minded	19.2	5.85
Dominant	15.9	3.65	Immature	28.0	5.77	Dominant	18.2	5.86
Generous	14.7	4.23	Irrational	27.8	4.81	Uncultured	17.7	6.12
Deceptive	13.6	3.60	Gentle	27.4	4.37	Cultured	16.4	5.44
Forgiving	12.0	5.02	Wise	26.0	4.58	Civilised	15.6	5.19
Gentle	11.3	4.19	Stupid	23.1	4.80	Submissive	13.9	5.40
Kind	10.4	4.00	Ethical	22.7	5.02	Intellectual	12.5	5.36
Capable	8.8	4.22	Modest	22.6	6.46	Inflexible	10.6	6.29
Honest	8.8	5.51	Controlling	21.7	5.41	Mature	8.6	6.04
Immature	8.7	4.43	Trusting	19.7	5.70	Capable	7.9	4.96
Error-prone	8.2	5.01	Genuine	18.8	5.86	Rational	7.6	5.87
Disciplined	7.6	4.38	Intellectual	13.8	5.63	Efficient	4.3	6.47
Helpful	7.5	4.54	Energetic	13.0	5.66	Sophisticated	2.4	5.77
Efficient	7.3	4.50	Unsophisticated	11.3	5.32	Passive	2.0	6.17
Cold	5.9	3.93	Superficial	9.0	5.90	Disciplined	1.2	6.10
Warm	5.8	4.15	Refined	7.3	5.32	Generous	1.2	5.78
Irrational	5.6	4.66	Error-prone	5.4	5.42	Modest	1.0	5.82
Aggressive	3.8	3.15	Honest	4.2	5.78	Refined	−0.3	6.30
Calm	3.6	3.36	Innocent	2.2	6.25	Moral	−1.0	6.45
Genuine	3.2	5.17	Sophisticated	1.1	6.05	Warm	−2.4	6.23
Impulsive	3.1	4.22	Unrefined	0.8	5.68	Knowledgeable	−4.2	6.41
Sneaky	0.8	4.03	Rational	0.4	5.89	Wise	−4.4	5.50
Inflexible	0.7	4.55	Knowledgeable	−0.2	4.94	Ethical	−5.5	6.61
Curious	0.6	3.57	Uncultured	−1.8	6.73	Innocent	−7.7	6.52
Stupid	0.6	5.14	Capable	−3.2	5.14	Humble	−9.4	5.79
Trusting	−0.9	4.73	Calm	−3.6	7.08	Helpful	−9.4	6.33
Energetic	−3.2	3.49	Helpful	−5.8	5.03	Trusting	−9.7	5.89
Passive	−4.7	3.64	Passive	−8.9	6.19	Gentle	−10.5	5.57
Unsophisticated	−5.0	5.91	Cold	−11.2	6.02	Genuine	−10.8	6.50
Submissive	−6.4	4.41	Submissive	−14.2	5.68	Forgiving	−11.8	6.17
Unrefined	−8.8	5.14	Disciplined	−14.4	5.77	Kind	−13.4	5.60
Uncultured	−11.9	6.28	Inflexible	−19.3	5.79	Honest	−16.8	6.29
Innocent	−13.3	5.22	Efficient	−21.6	3.90	Calm	−16.9	5.75

Mean (M) scores and standard error of the mean (SE) are presented alongside each word (+50 indicated the word only applies to humans, and −50 indicated the word only applies to the comparison category). A different yet overlapping set of qualities emerged depending on the comparison category. There were an approximately equal number of desirable and undesirable characteristics for the other species and robots comparisons, but almost all were undesirable for the angels comparisons.

2.1.3. Design and data presentation

Every participant rated each trait on the three comparison scales. We present the mean ratings for the sixty traits from most to least human on the three scales (Table 1).

2.1.4. Procedure

Participants were informed that the study aimed to help us understand the ways in which people ascribe character traits and would be asked to rate sixty trait words on three separate scales. Once informed consent was obtained, brief demographic (age and gender) and screening (English fluency) questions were asked. Then, participants were taken through the three question blocks, before completing the final brief attitude scales towards each of the comparison categories (other species, angels and robots) at the end. Participants were debriefed and redirected back to Prolific for payment. The study took approximately ten minutes.

2.2. Results

For each of the three scales, words were ranked in order of score from highest (+50, corresponding to *Just humans*) to lowest (−50, corresponding to *Just other species/robots/angels*). Means for each word on each scale are presented in Table 1. Distinct yet overlapping attributes emerged as being 'uniquely human' depending on the comparison category. When compared to other species and to robots, an approximately equal number of desirable and undesirable traits emerged. For example, both corrupt and cultured appear in the top five most human traits when compared to other species, while both jealous and generous appear in the top five most human traits when compared to robots. When compared to angels, unsurprisingly, all of the twenty most human traits were undesirable in valence, such as selfish, corrupt and aggressive, with the possible exception of cynical and cunning, which may be viewed as more ambiguous. This shows the importance of the comparison category when defining lay conceptions of humanness.

2.3. Pretest discussion

When developing the Dual Model, Haslam and colleagues (Haslam, 2006; Haslam et al., 2004; Haslam et al., 2005) found that people associate broadly positive attributes with humans including civility, moral sensibility and warmth. This could be because human cognition has been shown to be biased towards representing immoral events as impossible (Phillips & Cushman, 2017). It could also be explained by features of Haslam and colleagues' design, including which traits were included the stimulus set and the particular questions that were asked of participants.

In a pre-registered design, we asked participants to rate sixty traits on how human specific they perceived each to be on three separate scales: comparing humans to other species, robots and angels. In comparison to other species and robots, approximately half of the traits rated as specific to humans were undesirable and half were more desirable. Consistent with previous conceptualisations (Haslam, 2006; Haslam et al., 2005), participants viewed humans as more open-minded, civilised and sophisticated than other animals and more generous, warm and kind than robots. However, humans were also thought of as more corrupt, controlling and arrogant than other animals and more jealous, selfish and stingy than robots. When asked what distinguished humans from angels, the top twenty traits were almost all negative and socially undesirable.

The particular terms viewed as unique to humans varied depending on the comparison point. Traits thought of as unique to humans in the context of animals were somewhat different from the traits thought of as unique to humans in the context of robots and different traits again were thought of as unique to humans in the context of angels. This suggests apparent evidence for two senses of humanness may be a product of the questions asked. Haslam et al. (2005) asked their participants two questions and found evidence for two senses of humanness. When we asked three questions, we found evidence for three variations. Had we included more comparisons (for example, to different types of animal such as rats, lice, swans or horses, or to different types of machine such as medical robots, personal computers or drones), we may have found evidence for further variation still (Over, 2020a). Work from cognitive psychology demonstrating the importance of context in representations applies to our understanding of humanness as well (Medin & Smith, 1984; Smith & Medin, 1981; Yee & Thompson-Schill, 2016). A full investigation of which traits are most associated with humans is beyond the scope of this paper and we do not seek to provide an exhaustive account. Rather, we suggest that what counts as typical of humans varies with context and that current trait-based models of dehumanization are consequently incomplete.

Understanding lay conceptions of humanness as undesirable as well as desirable opens the question as to whether there would still be evidence for trait-based dehumanization of outgroups when we control for valence in measuring intergroup trait ascriptions. Using the uniquely human desirable and undesirable attributes identified in this pretest, in the next seven experiments, we compare the dual model of dehumanization with an alternative explanation – that previously reported effects primarily reflect ingroup preference.

3. Study 1: Measuring intergroup ascriptions of uniquely human traits

3.1. Experiment 1a: Testing animalistic dehumanization against ingroup favouritism in a political group context

In this experiment, we measure whether there is evidence for trait-based dehumanization of a political outgroup when we control for trait valence. Taking the top ten desirable words and the top ten undesirable words rated as most human compared to other species in our pretest, we asked participants to rate how typical each one is of ingroup and outgroup members. The dual model of dehumanization (Haslam, 2006) suggests that animalistic dehumanization occurs when uniquely human traits are assigned more strongly to the ingroup than the outgroup. In contrast, we predicted an interaction such that participants will tend to more strongly attribute desirable uniquely human traits to the ingroup, but undesirable uniquely human traits to the outgroup.

In the first instance, we tested these hypotheses with words rated uniquely human in the context of other species because animalistic dehumanization is the most commonly studied form of dehumanization in intergroup settings (Demoulin et al., 2004; Haslam, 2006; Leyens et al., 2000, 2001; Paladino et al., 2002; Smith, 2014; Viki et al., 2006). We utilised a political intergroup context - Brexiters vs. Remainers. This was a salient intergroup divide in the UK when the study was run (November/December 2019) in the lead-up to the December 2019 general election, when the outcome would affect whether the UK would leave the EU. We chose this pertinent intergroup division to maximise the chances of finding dehumanizing biases if they were prevalent in the population. Political group contexts feature widely in the dehumanization literature. Blatant instances include cases of visual and verbal propaganda portraying the opposition in animalistic caricatures and metaphors (e.g., discussed in Cassese, 2020). Lab-based evidence for more subtle dehumanization of political outgroups has also been reported (Pacilli et al., 2016). In addition, participants completed the blatant dehumanization scale (Kteily et al., 2015) in order to ensure that explicit dehumanization was observed in this context.

3.1.1. Methods

3.1.1.1. Participants. A sample size calculation indicated that 126 participants would be required to detect a medium effect size (partial eta^2 0.06) for the 2 × 2 interaction with power of 0.8 and alpha 0.05. We tested 130 participants based on this. To be eligible for the study,

participants had to be fluent in English and a UK resident for the intergroup divide to be meaningful. Seven additional participants failed one or more attention checks so their data were excluded and replaced following our pre-registered inclusion criteria. The final sample (80 female, 49 male, 1 non-binary) were aged between 18 and 76 (Mean age = 33.8, *SD* = 13.2), 116 were British, and 101 identified as Remainers and 29 as Brexiters.

3.1.1.2. Materials. We took the top ten desirable trait words and the top ten undesirable trait words most strongly rated as applying to humans compared to other species in the pretest. The desirable words were: cultured, open-minded, civilised, sophisticated, creative, knowledgeable, moral, refined, humble and wise. The undesirable words were: corrupt, controlling, arrogant, superficial, bitter, stingy, selfish, spiteful, jealous, cruel. Although 'cynical' and 'intellectual' were rated as highly human compared to other species, these were not included as they may be perceived as more ambiguous in valence. Table 2 shows the trait items included for all experiments in Study 1. When means for the perceived humanness of the desirable words (overall $M = 28.8 \pm 3.10$) and the undesirable words (overall $M = 28.2 \pm 2.51$) were created for each participant from the pretest, a paired samples *t*-test found that the desirable and undesirable words were viewed as equally characteristic of humans, $t(29) = 0.15$, $p = .883$, $d = 0.03$. This was supported by an estimated Bayes factor in favour of the null model, $BF_{01} = 5.01$ and ensured that humanness was adequately controlled for across the two valence levels.

Participants indicated the extent to which they thought each trait word was typical of Brexiters and of Remainers within two blocks (one for each group condition). For each item, participants indicated their response on a sliding scale from *Not at all* (0) to *Very much so* (100), with the midpoint *Somewhat* (50), though they could not see the actual numbers. For example, a block could begin 'In the following questions, please consider the group: Brexiters'. Then, participants would respond to each item, such as 'Brexiters are typically cultured'. The order of blocks was counterbalanced evenly across participants and the twenty items within block were randomised. One attention check per block was also included approximately halfway through, such as 'Please indicate *Not at all*'.

As well as the group attributions, participants completed group preference and blatant dehumanization scales. In the preference scale, participants were asked to indicate how they felt about each group (Brexit/Remain) using a sliding scale from *Extremely Negative* (0) to *Extremely Positive* (100), though again they could not see the numbers. In the 'blatant dehumanization' scale (Kteily et al., 2015), participants saw the 'ascent of man' image and were asked to indicate on a slider how evolved they considered the average member of each group to be, with 0 corresponding to the ape-like silhouette at the very bottom and 100 to the most 'human' at the very top. Again the numbers were not visible. The order of items (whether Brexit or Remain was first) was evenly counterbalanced for both scales.

3.1.1.3. Procedure. Participants were informed that the study was designed to help us understand the ways in which people ascribe character traits to different groups of individuals, in this case, Brexiters and Remainers and that it did not matter which group they supported in order to take part. They were instructed that they would be asked to rate twenty trait words on two scales, one with reference to how typical the trait is to Brexiters and the other with reference to how typical the trait is to Remainers and then would be asked to complete two scales asking about attitudes to each group. Once informed consent was obtained, brief demographic (age, gender, nationality, and whether they supported Brexit or Remain) and screening (English fluency, current country of residence) questions were asked. Then, participants were taken through the two group attribution question blocks. Following this, participants completed the group preference and then the blatant dehumanization scales. Lastly, participants were debriefed and redirected back to Prolific for payment. On average, the study took under seven minutes.

3.1.1.4. Design and planned data analysis. In line with our pre-registered analysis plan, Brexit and Remain group identities were collapsed across participants into simple ingroup and outgroup categories. We designed each of our studies in this way because we could not be confident in advance how many participants we would be able to recruit from each political group.

There were four conditions in total in a 2 × 2 within-subjects design. A 2 (target group: ingroup/outgroup) x 2 (valence: desirable/undesirable) within subjects ANOVA tested for the interaction between group membership and valence in trait attributions.

In all studies, significant interactions were followed up with planned comparisons measuring differences in ratings between ingroup and outgroup for each condition. Additionally, we measured differences in preference and 'blatant dehumanization' ratings between ingroup and outgroup using paired-samples *t*-tests. All data in the present work met the assumptions necessary for the parametric tests performed.

3.1.2. Results

3.1.2.1. Interaction between valence and group condition in trait attributions. A 2 (valence: desirable/undesirable) x 2(target group: ingroup/outgroup) within subjects ANOVA tested for an interaction between valence and group condition in trait attributions. Counter to evidence for trait-based dehumanization, there was no significant main effect of group, $F(1, 129) = 1.03$, $p = .313$, $\eta_p^2 = 0.008$, showing overall ratings to be of a similar magnitude for ingroup and outgroup. There was a significant main effect of valence, $F(1, 129) = 41.08$, $p < .001$, $\eta_p^2 = 0.242$, with higher ratings overall for desirable than for undesirable trait words. In line with ingroup favouritism effects, there was a significant

Table 2
Trait items included for each condition in Study 1.

Experiment 1a		Experiment 1b		Experiment 1c	
UH compared to other species		*UH compared to robots*		*UH compared to angels*	
Desirable	Undesirable	Desirable	Undesirable		
Cultured	Corrupt	Generous	Jealous	Selfish	Arrogant
Open-minded	Controlling	Open-minded	Selfish	Corrupt	Immature
Civilised	Arrogant	Warm	Stingy	Aggressive	Impulsive
Sophisticated	Superficial	Kind	Impulsive	Jealous	Deceptive
Creative	Bitter	Cultured	Bitter	Stingy	Stupid
Knowledgeable	Stingy	Moral	Spiteful	Cruel	Unsophisticated
Moral	Selfish	Creative	Arrogant	Error-prone	Cunning
Refined	Spiteful	Mature	Sneaky	Spiteful	Irrational
Humble	Jealous	Curious	Cruel	Bitter	Cold
Wise	Cruel	Forgiving	Aggressive	Cynical	Unrefined

Note: UH stands for uniquely human.

interaction between valence and group condition, $F(1, 129) = 226.98$, $p < .001$, $\eta_p^2 = 0.638$, such that ratings were higher for ingroup than outgroup on desirable words, but higher for outgroup than ingroup on undesirable words ($ps < 0.001$) (Fig. 1A).

3.1.2.2. Intergroup dehumanization and preference ratings. Paired samples t-tests found that participants gave significantly higher ratings for ingroup than outgroup both on the blatant dehumanization scale, $t(129) = 8.029$, $p < .001$, $d = 0.70$ and the group preference scale, $t(129) = 20.18$, $p < .001$, $d = 1.77$ (Fig. 2). This showed participants rated feeling more negatively towards the outgroup than the ingroup and also rated the outgroup as less evolved on the visual depiction of the 'ascent of man'.

3.1.3. Discussion

The dual model of dehumanization holds that when outgroups are animalistically dehumanized, they are typically thought to possess uniquely human attributes such as civility, refinement, moral sensibility, rationality, and maturity to a lesser extent than ingroups (Haslam, 2006). We tested whether there is evidence that an outgroup is animalistically dehumanized when undesirable as well as desirable uniquely human traits are considered. We found that participants rated desirable traits (such as cultured) as more typical of ingroup members than outgroup members, but undesirable traits (such as corrupt) as more typical of outgroup members than ingroup members. This was the case even though the desirable and undesirable traits were judged by a previous set of participants to be equally typical of humans. The results provide no evidence for animalistic dehumanization as characterised by the dual model in this group context. Rather, participants show evidence of ingroup preference, the tendency to ascribe more desirable characteristics to the ingroup than the outgroup (Hewstone, Rubin, & Willis, 2002; Macrae & Hewstone, 1990).

3.2. Experiment 1b - Testing mechanistic dehumanization against ingroup favouritism in a political group context

Here we extend results from Experiment 1a using another set of uniquely human words identified in our pretest. This time we included 20 words rated as characteristic of humans in the context of robots. In doing this, we tested for evidence of mechanistic dehumanization when typically human but undesirable words are incorporated into the stimulus set. According to the dual model, mechanistic dehumanization is proposed to correspond to a denial of human nature traits, which are those judged as characteristic of humans in a non-comparative sense. However, it is often noted that these are traits that distinguish humans from robots and other machines. Whilst Haslam and colleagues refer to these attributes as 'human nature', we call them uniquely human (compared to shared with robots) for consistency throughout the present work. As before, we took the top ten desirable and top ten undesirable words from those rated as most human, though this time in the context of robots. We tested for an intergroup bias in ascribing these terms by asking participants to rate how typical they believed them to be of ingroup and outgroup members. The dual model of dehumanization holds that in mechanistic dehumanization, traits that are most characteristic of humans should typically be assigned more strongly to the ingroup than outgroup. In contrast, we predict an interaction between trait valence and group membership such that people more strongly attribute desirable traits to the ingroup and undesirable ones to the outgroup. Group membership was again in the form of Brexiters and Remainers.

3.2.1. Methods

3.2.1.1. Participants. Based on the same sample size calculation as for Experiment 1a, we tested 130 participants. Eligibility criteria were the same as before and data from seven participants that failed one or more of the attention checks was excluded and replaced. The final sample (92 female, 36 male, 2 non-binary) were aged between 18 and 65 (Mean age = 33.2, $SD = 10.6$). 109 were British and 91 identified as Remainers and 39 as Brexiters.

3.2.1.2. Online survey. The questions were almost identical to those included in Experiment 1a but included the top ten desirable and top ten undesirable trait words most strongly rated as applying to humans compared to robots (rather than other species). The desirable words were: generous, open-minded, warm, kind, cultured, moral, creative, mature, curious, forgiving. The undesirable words were: jealous, selfish, stingy, impulsive, bitter, spiteful, arrogant, sneaky, cruel, aggressive. Some words that appeared in the top twenty in the pretest (emotional, cynical and cunning) were not included because of being more ambiguous in valence. Table 2 shows the trait items included for all experiments in Study 1. As before, a paired samples *t*-test found no differences between 'humanness' ratings for the desirable ($M = 34.4 \pm 2.18$) and undesirable ($M = 36.4 \pm 2.62$) words, $t(29) = 0.96$, $p = .345$, $d = 0.16$. This was again supported by an estimated Bayes factor in favour of the null model, $BF_{01} = 3.38$. The procedure, design and planned data analysis exactly followed Experiment 1a.

3.2.2. Results

3.2.2.1. Interaction between valence and group condition in trait attributions. There was no evidence for mechanistic dehumanization - there was no significant main effect of group, $F(1, 129) = 1.37$, $p = .244$, $\eta_p^2 = 0.010$, showing overall ratings to be of a similar magnitude for ingroup and outgroup. However, there was a significant main effect of valence, F

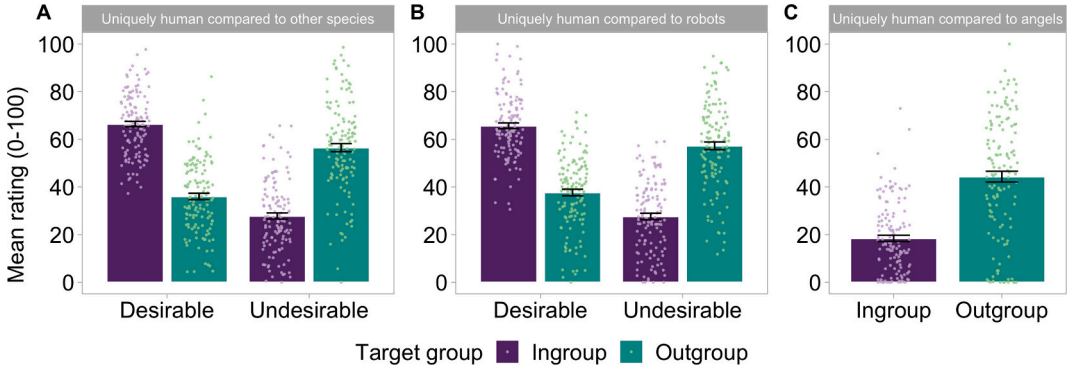

Fig. 1. Evidence for intergroup preference but not trait-based dehumanization in Experiments 1a (A), 1b (B) and 1c (C). Mean ratings were higher for political ingroup members than outgroup members on desirable uniquely human traits but higher for outgroup than ingroup members on undesirable uniquely human traits. Error bars represent standard errors.

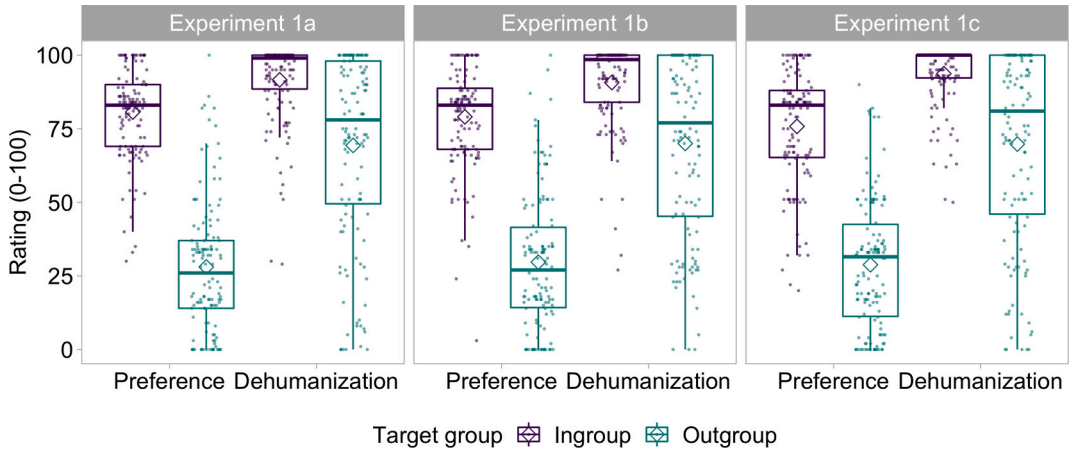

Fig. 2. Blatant dehumanization and preference scores for ingroup and outgroup in Experiments 1a, 1b and 1c. Participants rated the political outgroup as less 'human-like' than the ingroup across studies and all these differences were significant at $p < .001$.

(1, 129) = 53.16, $p < .001$, $\eta_p^2 = 0.292$, with higher ratings overall for desirable than for undesirable words. As we predicted, there was a significant interaction between valence and group condition, $F(1, 129) = 188.42$, $p < .001$, $\eta_p^2 = 0.594$. Ratings were higher for ingroup than outgroup on desirable words, but higher for outgroup than ingroup on undesirable words (*ps* <. 001) (Fig. 1B).

3.2.2.2. Differences in dehumanization and preference ratings for ingroup and outgroup. Participants gave higher ratings for ingroup than outgroup both on the blatant dehumanization scale (M ingroup = 90.7 ± 1.24; M outgroup = 70.0 ± 2.65), $t(129) = 8.27$, $p < .001$, $d = 0.72$ and on the group preference scale (M ingroup = 79.0 ± 1.49; M outgroup = 29.7 ± 2.03), $t(129) = 16.57$, $p < .001$, $d = 1.45$ (Fig. 2).

3.2.3. Discussion

This study aimed to extend findings from Experiment 1a using words viewed as uniquely human in the context of robots. Participants gave higher ratings for ingroup than outgroup members on desirable words, but higher ratings for outgroup than ingroup members on undesirable words. We found no evidence of mechanistic dehumanization in this context, but we did find evidence for ingroup favouritism. Once again, participants explicitly claimed that outgroup members seemed less human-like than ingroup members on the blatant dehumanization scale (Kteily et al., 2015), confirming this is the type of social context where we should see the preferential attribution of human traits to ingroups outlined by the dual model if they occur. Taken together, Experiments 1a and 1b suggest that political opponents in the Brexit v Remain group context were neither animalistically nor mechanistically dehumanized.

3.3. Experiment 1c - Testing dehumanization against ingroup favouritism with a third conception of humanness

In this study, we sought to extend the results of Experiments 1a and 1b to a third set of words rated as uniquely human. We took the top twenty words from our pretest that were rated as most strongly applying to humans compared to angels, which were generally undesirable in character. Though the dual model of dehumanization does not make specific predictions here, an extension of its logic implies that trait-based dehumanization, as a separable process from ingroup favouritism, would entail these human specific qualities to be ascribed more strongly to the ingroup than the outgroup. We again predicted effects of ingroup favouritism – even though these words are rated as human specific, they are likely to be attributed more strongly to the outgroup than the ingroup. As before, group membership was in the form of Brexiters and Remainers.

3.3.1. Methods

3.3.1.1. Participants. For consistency with the previous two studies, 130 participants were tested. Eligibility and exclusion criteria were the same as for Experiments 1a and 1b. Nine failed one or more attention checks so their data was excluded and replaced. The final sample (87 female, 41 male, 1 agender) were aged between 18 and 80 (Mean age = 36.0, $SD = 13.0$). 111 were British, 100 identified as Remainers and 30 as Brexiters.

3.3.1.2. Materials and design. The questions were almost identical to those included in Experiments 1a and 1b but this time we included the top twenty words that were most strongly rated as applying to humans compared to angels in the pretest. The words were: selfish, corrupt, aggressive, jealous, stingy, cruel, error-prone, spiteful, bitter, cynical, arrogant, immature, impulsive, deceptive, stupid, unsophisticated, cunning, irrational, cold, unrefined (Table 2). The procedure exactly followed that of Experiments 1a and 1b.

3.3.1.3. Design and planned data analysis. The design and planned data analysis was similar to the previous experiments but this time there were just two conditions, ingroup and outgroup target. A paired-samples *t*-test measured whether participants differentially attributed these undesirable characteristics to the two groups. The same analyses as in the previous studies examined differences in group preferences and dehumanization ratings.

3.3.2. Results

3.3.2.1. Main effect of group condition on (undesirable) trait attributions. Overall ratings were significantly higher for the outgroup ($M = 44.3 \pm 2.26$) than the ingroup ($M = 18.4 \pm 1.30$), $t(129) = 12.09$, $p < .001$, $d = 1.06$ (Fig. 1C). This showed participants tended to indicate that undesirable traits, despite being high in 'humanness', were more typical of the outgroup than the ingroup.

3.3.2.2. Differences in dehumanization and preference ratings for ingroup and outgroup. Once again, ratings were significantly higher for ingroup than outgroup both on the blatant dehumanization scale, $t(129) = 8.86$, $p < .001$, $d = 0.78$ and on the preference scale, $t(129) = 15.02$, $p < .001$, $d = 1.32$ (Fig. 2).

3.3.3. Discussion

This experiment demonstrated that participants rated undesirable but uniquely human words as more typical of outgroup members than

ingroup members. The results support our view that apparent evidence for trait-based dehumanization can be better explained by ingroup preference. Though these results are not intuitively surprising, they highlight a need for fully separating effects of valence from effects of humanness in intergroup bias in trait attributions. Dovetailing with recent critiques of the dual model, we show that outgroups are perceived in uniquely human terms, though socially undesirable ones (Over, 2020a; Over, 2020b). Previous evidence for trait-based dehumanization may at least in part be obscured by a conflation of 'humanness' with 'good'. As before, participants rated outgroup members as lower on the blatant dehumanization scale (Kteily et al., 2015) showing that this is the type of intergroup division where the trait-based accounts of dehumanization ought to predict preferential attribution of human traits to ingroup members.

Taken together, these three experiments bring into doubt the argument that trait-based dehumanization (Haslam, 2006), is a distinct cognitive process from intergroup preference. In the next four experiments, we test this more thoroughly by replicating the experiments in two further intergroup contexts and also including non-uniquely human traits (those shared with other species and robots) into the stimulus sets.

4. Study 2: Measuring intergroup ascriptions of uniquely human and non-uniquely human traits (immigrant outgroup)

4.1. Experiment 2a: Testing animalistic dehumanization against ingroup favouritism in the context of immigration

In Study 2, we seek to replicate and extend our findings in another intergroup context. In this study, we measure whether there is evidence for trait-based dehumanization of the two types outlined by the dual model in the context of immigrants to the UK. We chose this intergroup context because of its social significance. Immigrant groups face systematic biases in multiple contexts, from discrimination by the police to employer hiring decisions, impacting important outcomes such as those relating to health and education (Fernández-Reino, 2019, 2020; Kauff, Wölfer, & Hewstone, 2017). Furthermore, previous research has suggested that immigrants are often dehumanized, both blatantly, for example on the explicit dehumanization scale (Kteily et al., 2015; Kteily & Bruneau, 2017; Markowitz & Slovic, 2020), and more subtly, by being denied human specific mental states (Banton, West, & Kinney, 2020).

In addition to replicating our findings, we also add a further manipulation in which we compare attribution of traits that vary in humanness (uniquely human or non-uniquely human) as well as in valence. We incorporated this factor so we could more thoroughly separate potential dehumanization effects as predicted by the dual model from effects of valence. It is possible that trait-based dehumanization could be evidenced in an interaction between target group and humanness (with uniquely human words, but not non-uniquely human words ascribed more strongly overall to the ingroup than the outgroup, for example). Evidence for dehumanization of this sort could exist alongside, but independently of, valence effects. This design is also more consistent with prior related work (Hodson & Costello, 2007; Leyens et al., 2001; Viki et al., 2006).

In Experiment 2a, we test whether there is evidence for animalistic dehumanization of immigrants when the valence of the trait terms used is controlled for. We compare how British participants attribute uniquely and non-uniquely human characteristics that vary orthogonally in desirability (desirable/undesirable) differentially to immigrants and UK nationals who reside in the UK. The dual model predicts that in cases of animalistic dehumanization, uniquely human traits will overall typically be attributed more strongly to the ingroup than the outgroup. This should not be true of traits shared with other species. These interaction effects should be independent of trait valence. However, we predict that desirable traits will typically be attributed more strongly to the ingroup than outgroup and undesirable traits more strongly to the outgroup than the ingroup, regardless of perceived humanness.

4.1.1. Methods

4.1.1.1. Participants. A sample size calculation indicated that 126 participants would be required to detect a medium effect size (partial eta^2 0.06) for the 2x2x2 interactions with power of 0.8 and alpha 0.05. We tested 130 participants. Two participants failed one or more attention checks and were excluded and replaced. In order to maximise our chances of finding evidence of trait based dehumanization if it occurs, we chose to only test people that self-identified as right-wing and that indicated they voted for Brexit in the 2016 referendum. This is because previous research has demonstrated that right-wing participants are particularly liked to dehumanize immigrants (Markowitz & Slovic, 2020). Further, anti-immigration attitudes were widespread amongst Brexit supporters (Goodwin & Milazzo, 2017; Meleady, Seger, & Vermue, 2017). Participants were eligible if they were fluent in English, UK nationals and residents, voted Brexit, and indicated they affiliate to the right of the political spectrum. The final sample (61 female, 69 male) were aged between 22 and 78 (Mean age = 45.6, SD = 14.3).

4.1.1.2. Materials. To create the stimuli, we chose trait words from our pretest data (Table 1) that best fit our four categories of interest: unique to humans and desirable, unique to humans and undesirable, shared with other species and desirable, and shared with other species and undesirable. To ensure our conceptions of trait desirability were accurate, we separately asked thirty different participants to rate each of the sixty traits on a desirability scale. This asked participants to 'indicate the extent to which you think the word in each of the following questions is a desirable character trait for someone to have.' The bottom end of the slider (0) indicated the trait to be highly undesirable (a negative/bad quality to have), while the top end (100) indicated the trait to be highly desirable (a positive/good quality to have). Full results from these ratings can be found in the OSF page (https://osf.io/yp8wc/) provided for this work.

From both the most and least uniquely human traits compared to other species, we chose five that were desirable and five that were undesirable. Of the uniquely human category, the desirable traits were cultured, open-minded, civilised, sophisticated and knowledgeable, and undesirable traits were corrupt, controlling, arrogant, superficial and bitter. Of the traits shared with other species (least uniquely human), the desirable traits were energetic, trusting, curious, genuine and calm, and the undesirable traits were uncultured, unrefined, unsophisticated, stupid and inflexible (Table 3 shows the trait words included within each condition). In support of our experimental manipulations, paired samples *t*-tests showed that the mean of the uniquely human traits were rated as significantly more human than of those shared with other species, both for the desirable condition, $t(29) = 9.44$, $p < .001$, $d = 1.72$, and for the undesirable condition, $t(29) = 8.88$, $p < .001$, $d = 1.62$. Additionally, the desirable traits were rated as significantly more desirable than the undesirable traits, both for the uniquely human

Table 3
Trait items included for each condition in Studies 2 & 3.

	Experiments 2a & 3a		Experiments 2b & 3b	
	Uniquely human	*Shared, other species*	*Uniquely human*	*Shared, robots*
Desirable	Cultured	Energetic	Generous	Efficient
	Open-minded	Trusting	Open-minded	Disciplined
	Civilised	Curious	Warm	Helpful
	Sophisticated	Genuine	Kind	Calm
	Knowledgeable	Calm	Moral	Capable
Undesirable	Corrupt	Uncultured	Jealous	Inflexible
	Controlling	Unrefined	Selfish	Submissive
	Arrogant	Unsophisticated	Stingy	Cold
	Superficial	Stupid	Impulsive	Passive
	Bitter	Inflexible	Bitter	Uncultured

condition, $t(29) = 20.33$, $p < .001$, $d = 3.71$, and for the non-uniquely human condition, $t(29) = 19.13$, $p < .001$, $d = 3.49$.

We chose the items for our trait categories such that humanness ratings did not significantly differ between the desirable and undesirable conditions for each level of humanness. Importantly, humanness scores were comparable for desirable and undesirable traits that were unique to humans, $t(29) = 0.002$, $p = .999$, $d < 0.001$, $BF_{01} = 5.14$) and for desirable and undesirable traits shared with other species $t(29) = 1.12$, $p = .274$, $d = 0.20$, $BF_{01} = 2.92$. Also, desirability scores were comparable for uniquely human and non-uniquely human traits that were desirable, $t(29) = 1.34$, $p = .192$, $d = 0.24$, $BF_{01} = 2.30$, and undesirable, $t(29) = 1.14$, $p = .264$, $d = 0.21$, $BF_{01} = 2.86$. This ensured that dimensions of Valence and Humanness were orthogonal, allowing us to accurately separate effects of each.

The scales were almost identical to those described for Study 1. Participants indicated the extent to which they thought each trait word was typical of UK nationals (ingroup) and of immigrants (outgroup) within two blocks (one for each target group condition). For each item, participants indicated their response on a sliding scale from *Not at all* (0) to *Very much so* (100), with the midpoint *Somewhat* (50), though they could not see the actual numbers. For example, a block could begin 'In the following questions, please consider the group: immigrants (to the UK from overseas)'. Then, participants would respond to each item, such as 'Immigrants are typically corrupt'. The order of blocks was counterbalanced evenly across participants such that half responded to the ingroup block first and half responded to the outgroup block first. The twenty items within block were randomised. One attention check per block was included approximately halfway through. Again, participants also completed the group preference and blatant dehumanization scales (Kteily et al., 2015). For both these attitude scales, the order of items (whether participants' ingroup or outgroup was first) was counterbalanced and accorded with order of presentation for the trait attribution blocks.

4.1.1.3. Procedure. The procedure took an almost identical format to those outlined for Study 1. Participants were informed that the study was designed to help us understand the ways in which people ascribe character traits to different groups of individuals, UK nationals and immigrants. They were instructed that they would be asked to rate twenty trait words on two scales, one with reference to how typical the trait is of UK nationals and the other with reference to how typical the trait is to immigrants and then would be asked to complete two scales asking about attitudes towards each group. Informed consent, brief demographics (age and gender) and screening (English fluency, political party identification, nationality and country of residence) questions were asked before participants were taken through the trait attribution questions. Participants then completed the group preference and blatant dehumanization scales before being debriefed and redirected for payment.

4.1.1.4. Design and planned data analysis. There were three within-subjects variables each with two levels (target group membership: ingroup/outgroup; trait humanness: uniquely human/shared with other species; trait valence: desirable/undesirable). In line with our pre-registered analysis plan, we conducted a 2 (target group: ingroup/outgroup) x2 (trait humanness: uniquely human/shared with other species) x2 (trait valence: desirable/undesirable) within-subjects ANOVA to test for interactions between target group membership, valence and humanness in intergroup trait attributions.

Significant interactions were followed up with planned analyses of simple effects. As for all reported experiments, these always focussed only on differences in ratings between ingroup and outgroup targets, as central to our pre-registered hypotheses. We measured differences in preference and 'blatant dehumanization' ratings between ingroup and outgroup using paired-samples *t*-tests.

4.1.2. Results

Scores for each emotion category were obtained by calculating the mean of the five trait terms within each category for each participant. For example, a participant's score for uniquely human desirable trait ascriptions towards the ingroup would be the mean of their ratings on cultured, open-minded, civilised, sophisticated and knowledgeable on the ingroup block.

4.1.2.1. Main ANOVA: Interaction between target group, trait humanness and trait valence. A 2(target group: ingroup/outgroup) x 2 (humanness: uniquely human/shared with other species) x 2 (valence: desirable/undesirable) within-subjects ANOVA tested for the interactions of interest. There were significant two-way interactions between group and valence, $F(1, 129) = 27.86$, $p < .001$, $\eta p^2 = 0.178$, group and humanness, $F(1, 129) = 10.23$, $p = .002$, $\eta p^2 = 0.073$, and valence and humanness, $F(1, 129) = 11.24$, $p = .001$, $\eta p^2 = 0.080$. The three-way interaction between group, humanness and valence was not significant, $F(1, 129) = 0.90$, $p = .346$, $\eta p^2 = 0.007$. Planned simple effects comparisons following up the Group x Valence interaction showed that ratings were higher for ingroup than outgroup on desirable traits ($p < .001$) but higher for outgroup than ingroup on undesirable traits ($p = .005$). Simple effects comparisons following up the Group x Humanness interaction showed that ratings were higher for ingroup than outgroup on both uniquely human traits ($p < .001$) and those shared with other species ($p = .002$).

Though the three-way interaction was not significant, we report the planned comparisons between ingroup and outgroup on each emotion condition so as to thoroughly assess both predictions and for consistency across experiments. Ratings were higher for ingroup than outgroup on desirable terms, both uniquely human and shared with other species ($ps < 0.001$), and higher for outgroup than ingroup on undesirable terms, both uniquely human ($p = .037$) and shared with other species ($p = .001$) (Fig. 3A).

4.1.2.2. Dehumanization and attitudes measures. Paired *t*-tests showed that participants indicated greater negative feeling to the outgroup than the ingroup, $t(129) = 10.76$, $p < .001$, $d = 0.94$. Participants also rated the outgroup as 'less human' than the ingroup on the blatant dehumanization scale, $t(129) = 4.80$, $p < .001$, $d = 0.42$ (Fig. 4).

4.1.3. Discussion

We tested for trait-based animalistic dehumanization of immigrants to the UK by right-wing Brexit supporters. In doing so, we replicated and extended results of Study 1 by including a further intergroup context. We also incorporated non-uniquely human as well as uniquely human trait words into our stimulus set, allowing us to measure intergroup attributions of traits that varied orthogonally on humanness and desirability. We did not observe the pattern of results that would demonstrate animalistic dehumanization by the dual model account. We found that immigrants were ascribed desirable traits to a lesser extent but undesirable traits to a greater extent than the ingroup, regardless of trait humanness. We replicated the pattern of results observed in Study 1 and showed again that intergroup biases in trait attribution are driven by valence rather than by humanness. Next, we tested for effects of mechanistic dehumanization in the same group context.

4.2. Experiment 2b: Testing mechanistic dehumanization against ingroup favouritism in the context of immigration

In Experiment 2b, we test whether there is evidence that immigrants are mechanistically dehumanized. We compared the extent to which participants attributed traits that varied on human uniqueness (compared to robots) and on desirability to immigrants and to UK nationals living in the UK. The dual model predicts that in cases of mechanistic dehumanization, qualities that are characteristic of humans

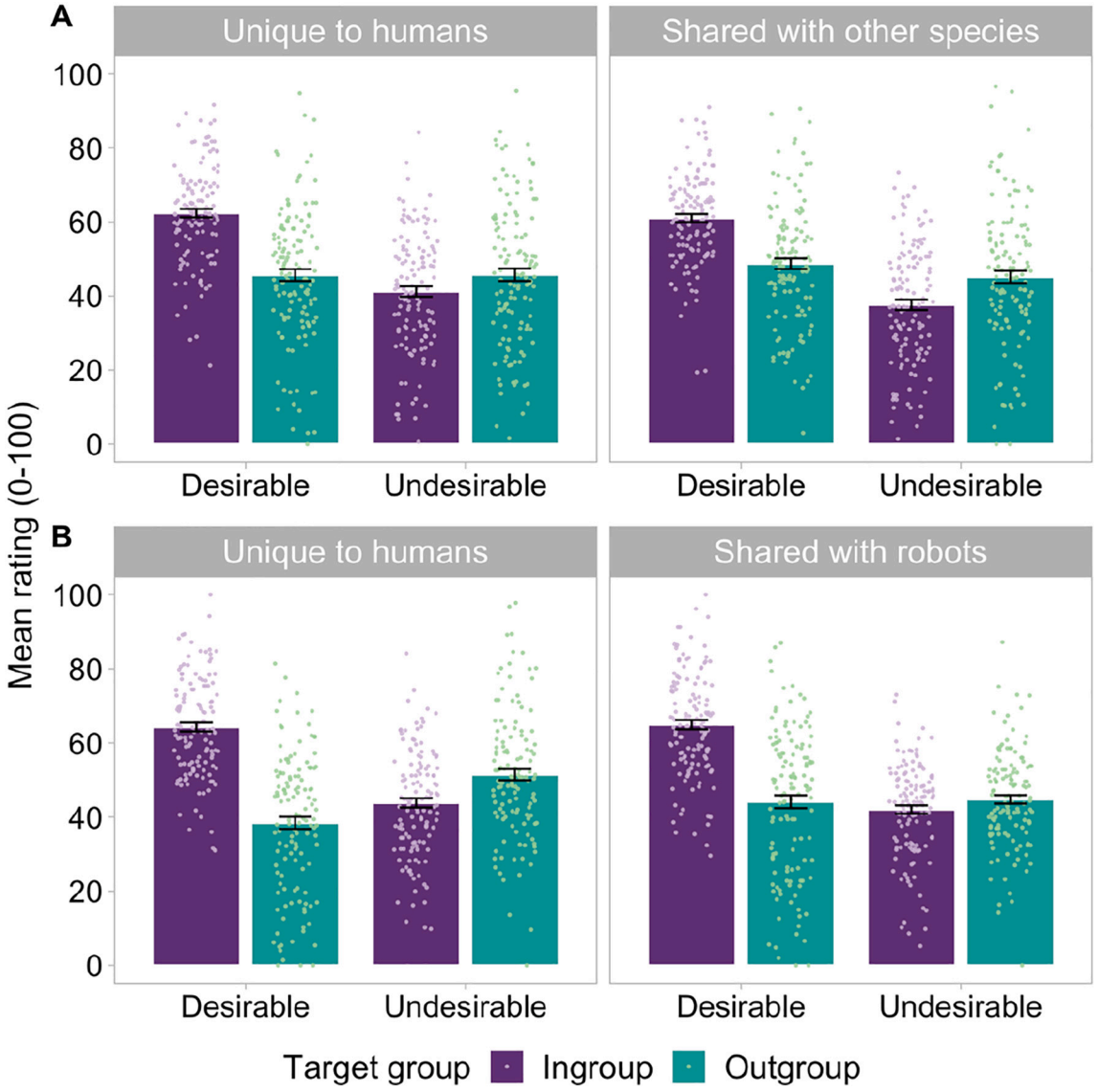

Fig. 3. Evidence for intergroup preference but not trait-based dehumanization in Experiments 2a (A) and 2b (B). Mean ratings were higher for the ingroup than for immigrants (outgroup) on desirable traits but higher for immigrants than for ingroup members on undesirable traits. This held both for traits perceived as uniquely human, and for those shared with other species/robots. Error bars represent standard errors.

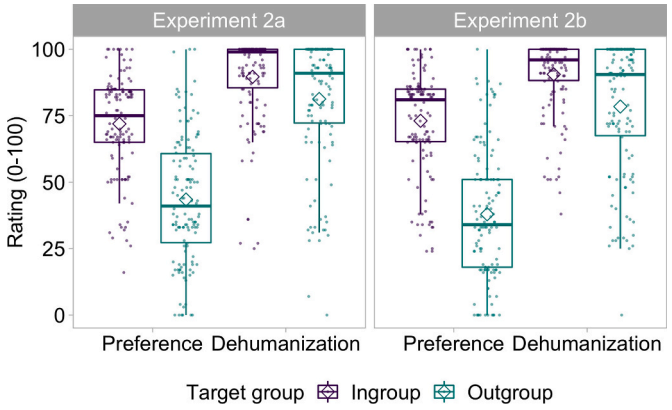

Fig. 4. Blatant dehumanization and preference scores for ingroup and outgroup in Experiments 2a and 2b. Participants rated immigrants as less 'human-like' than the ingroup across studies and all these differences were significant at $p < .001$.

(akin to those that distinguish humans from robots) will overall typically be attributed more strongly to the ingroup than to the outgroup. This should not be the case for traits shared with robots. These interaction effects should be independent of trait valence. However, we predict that desirable traits will typically be attributed more strongly to the ingroup than outgroup and undesirable traits more strongly to the outgroup than the ingroup, regardless of perceived humanness. If what appears to be evidence for trait-based dehumanization in previous research can better be described in terms of ingroup preference and stereotyping, then immigrants will tend to be attributed desirable traits less strongly and undesirable traits more strongly regardless of whether they are distinctly human traits or not.

4.2.1. Methods

4.2.1.1. Participants. As before, 130 participants were included. Eligibility and exclusions criteria were the same as for Experiment 2a. Three people failed one or more attention checks and their data was excluded and replaced. The final sample (56 female, 74 male) were aged between 22 and 80 (Mean age = 45.4, SD = 15.0).

4.2.1.2.

The scale items were developed using a similar approach as for Experiment 2a. However, this time, instead of including words as being most and least unique to humans compared to other species, we included words that were most or least unique to humans compared to robots. From our pretest data (Table 1), we again chose trait words to represent the four categories of interest: unique to humans and desirable, unique to humans and undesirable, shared with robots and desirable, and shared with robots and undesirable.

From both the most and least uniquely human terms compared to robots, we chose five that were desirable and five that were undesirable. Of the uniquely human category, the desirable traits were generous, open-minded, warm, kind and moral, and the undesirable traits were jealous, selfish, stingy, impulsive and bitter. Of the traits shared with robots (non-uniquely human), the desirable traits were efficient, disciplined, helpful, calm and capable, and the undesirable traits were inflexible, submissive, cold, passive and uncultured (Table 3 shows the traits included within each condition). In support of our experimental manipulations, paired *t*-tests showed that the combined means of the uniquely human traits were rated as significantly more human than of those shared with robots, both for the desirable, $t(29) = 9.97$, $p < .001$, $d = 1.82$, and for the undesirable, $t(29) = 10.99$, $p < .001$, $d = 2.00$, condition. Additionally, the desirable traits were rated as significantly more desirable than the undesirable traits, both for the uniquely human, $t(29) = 16.62$, $p < .001$, $d = 3.03$, and non-uniquely human, $t(29) = 17.53$, $p < .001$, $d = 3.20$, condition.

We again chose the items for our trait categories such that humanness and valence could be measured as orthogonal factors. Humanness scores were comparable for desirable and undesirable traits that were unique to humans, $t(29) = 1.02$, $p = .315$, $d = 0.19$, $BF_{01} = 3.19$, and for desirable and undesirable traits shared with robots, $t(29) = 0.42$, $p = .676$, $d = 0.08$, $BF_{01} = 4.74$. Desirability scores were comparable for uniquely human and non-uniquely human traits that were desirable, $t(29) = 1.46$, $p = .155$, $d = 0.27$, $BF_{01} = 1.98$, and undesirable, $t(29) = 0.51$, $p = .612$, $d = 0.09$, $BF_{01} = 4.55$.

Other than including different trait items, the scales were identical as described for Experiment 2a. The procedure, design and planned data analysis were also identical as for Experiment 2a.

4.2.2. Results

4.2.2.1. Main ANOVA: Interaction between target group, trait humanness and trait valence. A 2(target group: ingroup/outgroup) x 2 (humanness: uniquely human/shared with robots) x 2 (valence: desirable/undesirable) within-subjects ANOVA tested for the interactions of interest.

There were significant two-way interactions between group and valence, $F(1, 129) = 69.22$ $p < .001$, $\eta_p^2 = 0.349$, valence and humanness, $F(1, 129) = 50.39$, $p < .001$, $\eta_p^2 = 0.281$, but not between group and humanness, $F(1, 129) = 0.08$, $p = .936$, $\eta_p^2 < 001$. Effects were qualified in a significant three-way interaction between group, humanness and valence, $F(1, 129) = 21.40$, $p < .002$, $\eta_p^2 = 0.142$. Planned analyses of simple effects following the three-way interaction showed that ratings were higher for ingroup than outgroup on desirable terms, both uniquely human and shared with robots ($ps < 0.001$), and higher for outgroup than ingroup on uniquely human undesirable terms ($p < .001$). Although ratings were slightly higher for outgroup than ingroup on undesirable terms shared with robots, this did not reach significance ($p = .085$).

4.2.2.2. Dehumanization and attitudes measures. Paired t-tests showed that participants indicated greater negative feeling to the outgroup than the ingroup, $t(129) = 13.25$, $p < .001$, $d = 1.16$. Participants also rated the outgroup as less 'human' than the ingroup on the blatant dehumanization scale, $t(129) = 6.42$, $p < .001$, $d = 0.56$ (Fig. 4).

4.2.3. Discussion

Experiment 2b extended results from Experiment 2a and showed no evidence for trait-based mechanistic dehumanization of immigrants. Instead, we found that immigrants were ascribed desirable traits to a lesser extent than ingroup members, regardless of perceived humanness. Immigrants were also ascribed undesirable human specific traits to a greater extent than the ingroup. However, there was only a marginal difference between immigrants and the ingroup for undesirable traits shared with robots. This last results may reflect the importance of stereotypes and specific social context as well as ingroup preference effects in trait attributions (e.g., Fiske, Cuddy, Glick, & Xu, 2002).

Taken together, the experiments suggest that trait-based dehumanization does not occur in this context when trait valence is adequately controlled for. Rather, the observed pattern of results across Study 2 suggests that immigrants tend to be attributed desirable qualities less strongly and undesirable qualities more strongly than ingroup members.

One possible explanation for these results is that immigrants tend not to be dehumanized and, as a result, we did not observe the relevant differences in trait attribution. However, this explanation seems unlikely. Prior research has reported dehumanization of immigrants in multiple contexts (Banton et al., 2020; Kteily & Bruneau, 2017; Markowitz & Slovic, 2020), showing this is the kind of group we should expect to see dehumanized should the process occur. We also chose a sample of self-identified right-wing Brexit supporters to maximise chances of detect dehumanization effects towards immigrants (Markowitz & Slovic, 2020). Further, we found that immigrants were explicitly dehumanized on the blatant dehumanization scale by this sample. Our results show that trait-based accounts do not accurately characterise the nature of dehumanization when trait valence is appropriately controlled. Taken together, results from Study 2 hold important real-world implications. Immigrants face systematic discrimination that is directly linked to negative life outcomes (Fernández-Reino, 2019, 2020; Kauff et al., 2017). It is important for society to accurately understand the mechanisms underlying discrimination of this sort.

5. Study 3: Measuring intergroup ascriptions of uniquely human and non-uniquely human traits (criminal outgroup)

5.1. Experiment 3a: Testing animalistic dehumanization against ingroup favouritism in the context of criminals

In Study 3, we seek to replicate the findings from Study 2 in a third intergroup context. We do this in order to demonstrate the overarching challenge to the dual model and to test whether our alternative view generalises across multiple contexts.

In this study, we measure whether there is evidence that criminals are animalistically and/or mechanistically dehumanized when trait desirability is controlled for. We chose criminals as the outgroup because previous research utilising the dual model framework has suggested criminals are dehumanized along both dimensions. These studies show that the extent to which criminals are dehumanized has implications for sentencing judgements and support for rehabilitation programmes (Bastian et al., 2013; Viki et al., 2012). Therefore, understanding the specific mechanisms for biases in trait judgements towards criminals has important consequences within the criminal justice system and beyond. This intergroup context also meant we could widen the sample of participants belonging to the 'ingroup'. This allowed us to confirm prior results were not limited by the specific samples included (UK Brexit/Remain supporters in Study 1, and UK right wing Brexiters in Study 2).

The methods, analyses and predictions were the same as for Study 2. Animalistic dehumanization by the dual model account would be evidenced in uniquely human traits being attributed less strongly to criminals than to the ingroup. However, this should not be the case for traits shared with animals. This effect of humanness should be independent of trait valence. However, as for prior studies, we predict that desirable

traits will typically be attributed more strongly to the ingroup than to criminals and undesirable traits more strongly to criminals than the ingroup, regardless of humanness.

5.1.1. Methods

5.1.1.1. Participants. As for previous studies, 130 participants were included. Participants were eligible to take part if they were fluent in English, eighteen or over and had not previously served a prison sentence for committing a crime. Three participants failed one or more attention checks and their data was excluded and replaced. Participants were of a range of nationalities and residencies. The final sample (59 female, 70 male, 1 'other') were aged between 18 and 50 (Mean age = 25.4, SD = 7.3).

5.1.1.2. Materials, procedure and design. Materials, procedure and design were exactly the same as for Experiment 2a though this time participants indicated the extent to which they thought each trait word was typical of 'individuals with no criminal history' (ingroup) and of 'individuals with criminal convictions' (outgroup). For example, a block could begin 'In the following questions, please consider the group: individuals with criminal convictions'. Then, participants would respond to each item, such as 'individuals with criminal convictions are typically sophisticated'. As before, the order of blocks was counterbalanced evenly across participants.

5.1.2. Results

5.1.2.1. Main ANOVA: Interaction between target group, trait humanness and trait valence. A 2(target group: ingroup/outgroup) x 2 (humanness: uniquely human/shared with other species) x 2 (valence: desirable/undesirable) within-subjects ANOVA tested for the interactions of interest.

There were significant two-way interactions between group and valence, $F(1, 129) = 143.20$, $p < .001$, $\eta_p^2 = 0.526$ and valence and humanness, $F(1, 129) = 60.39$, $p < .001$, $\eta_p^2 = 0.319$, but not between group and humanness, $F(1, 129) = 2.20$, $p = .140$, $\eta_p^2 = 0.017$. Effects were qualified in a significant three-way interaction, $F(1, 129) = 9.18$, $p = .003$, $\eta_p^2 = 0.066$. Planned analyses of simple effects showed that ratings were significantly greater for ingroup than outgroup on desirable traits, both uniquely human and shared with other species, but significantly greater for outgroup than ingroup on undesirable traits, both uniquely human and shared with other species (all $ps < 0.001$) (Fig. 5A).

5.1.2.2. Dehumanization and attitudes measures. Paired *t*-tests showed that participants rated feeling more negatively towards the outgroup than the ingroup $t(129) = 16.92$, $p < .001$, $d = 1.48$, and also rated the outgroup as 'less human' than the ingroup on the blatant dehumanization scale, $t(129) = 7.02$, $p < .001$, $d = 0.62$ (Fig. 6).

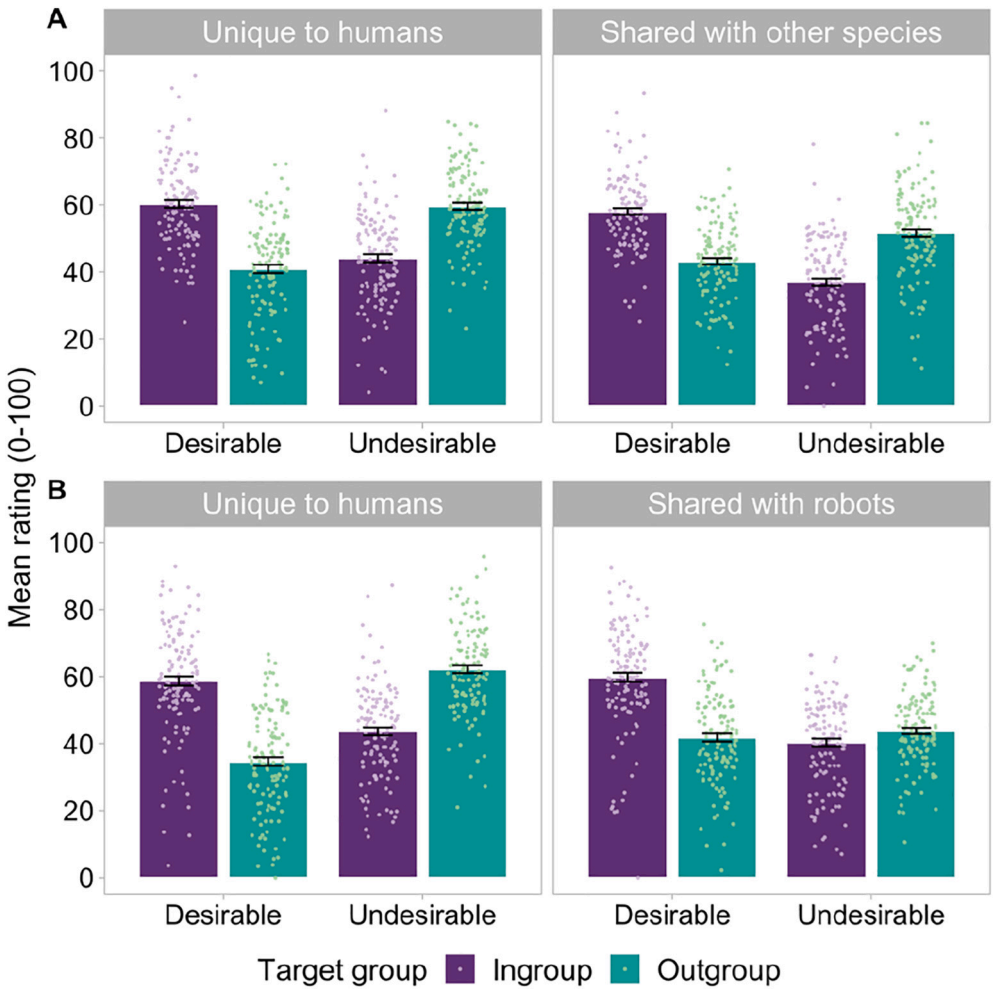

Fig. 5. Evidence for intergroup preference but not trait-based dehumanization in Experiments 3a (A) and 3b (B). Mean ratings were higher for the ingroup than for criminals (outgroup) on desirable traits but higher for criminals than for ingroup members on undesirable traits. This held both for traits perceived as uniquely human, and for those shared with other species/robots. Error bars represent standard errors.

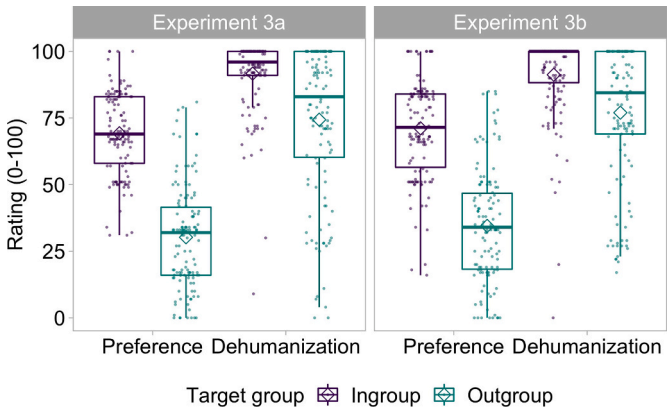

Fig. 6. Blatant dehumanization and preference scores for ingroup and outgroup in Experiments 3a and 3b. Participants rated criminals as less 'human-like' than the ingroup across studies and all these differences were significant at $p < .001$.

5.1.3. Discussion

We tested for trait-based animalistic dehumanization of criminals by the general population. This directly replicated Experiment 2a in a different intergroup context. Again, we did not observe the pattern of results that would demonstrate animalistic dehumanization by the dual model account. Rather, criminals were consistently ascribed desirable traits to a lesser extent but undesirable traits to a greater extent than the ingroup, irrespective of trait humanness. These results are inconsistent with prior work reporting animalistic dehumanization of criminals (Bastian et al., 2013; Viki et al., 2012). Consistent with findings from Study 2, we showed that intergroup biases in trait attribution in the context of criminals are driven by valence rather than by humanness. Next, we tested for effects of mechanistic dehumanization in the same intergroup context.

5.2. Experiment 3b: Testing mechanistic dehumanization against ingroup favouritism in the context of criminals

In Experiment 3b, we extend results by measuring whether there is evidence that criminals are mechanistically dehumanized when trait desirability is controlled for. This experiment replicates Experiment 2b exactly other than including criminals as the outgroup and non-criminals as the ingroup. As before, mechanistic dehumanization of criminals would be shown in weaker attributions of qualities that are characteristic of humans (akin to those that distinguish humans from robots) relative to ingroup members. This should not be the case for qualities that are shared with robots. These effects should be independent of valence. However, we predict that desirable traits will typically be attributed more strongly to the ingroup than to criminals and undesirable traits more strongly to the criminals than the ingroup, regardless of perceived humanness.

5.2.1. Methods

5.2.1.1. Participants. As for previous studies, 130 participants were included. Participants were eligible to take part if they were fluent in English, eighteen or over and had not previously served a prison sentence for committing a crime. Nine participants failed one or more attention checks and their data was excluded and replaced. Participants were of a range of nationalities and residencies. The final sample (71 female, 57 male, 1 'other' and 1 'prefer not to say') were aged between 18 and 69 (Mean age = 29.9, SD = 11.4).

5.2.1.2. Materials, procedure and design. Materials, procedure and design were exactly the same as for Experiment 2b though this time participants indicated the extent to which they thought each trait word was typical of 'individuals with no criminal history' (ingroup) and of 'individuals with criminal convictions' (outgroup), as for Experiment 3a.

5.2.2. Results

5.2.2.1. Main ANOVA: Interaction between target group, trait humanness and trait valence. A 2(target group: ingroup/outgroup) x 2 (humanness: uniquely human/shared with robots) x 2 (valence: desirable/undesirable) within-subjects ANOVA tested for the interactions of interest. There were significant interactions between group and humanness, $F(1, 129) = 19.19$, $p < .001$, $\eta_p^2 = 0.129$, group and valence, $F(1, 129) = 136.55$, $p < .001$, $\eta_p^2 = 0.514$ and valence and humanness, $F(1, 129) = 157.34$, $p < .001$, $\eta_p^2 = 0.549$. Effects were qualified by a significant three-way interaction, $F(1, 129) = 104.50$, $p < .001$, $\eta_p^2 = 0.448$. Planned simple effects analyses showed that ratings were significantly greater for ingroup than outgroup on desirable traits, both uniquely human and shared with robots ($ps < 0.001$), but significantly greater for outgroup than ingroup on undesirable traits, both uniquely human ($p < .001$) and shared with robots ($p = .016$) (Fig. 5B).

5.2.2.2. Dehumanization and attitudes measures. Paired *t*-tests showed that participants rated feeling more negatively towards the outgroup than the ingroup $t(129) = 13.78$, $p < .001$, $d = 1.21$ and also rated the outgroup 'less human' than the ingroup on the blatant dehumanization scale, $t(129) = 6.84$, $p < .001$, $d = 0.60$ (Fig. 6).

5.2.3. Discussion

Consistent with findings from Studies 1 and 2, we found no evidence for trait-based mechanistic dehumanization of criminals. Instead, criminals were ascribed desirable traits to a lesser extent than the ingroup and undesirable traits to a greater extent than the ingroup. These effects held for both uniquely human traits and for those perceived as shared with robots. Our results suggest that when desirability of the human specific attributes is rigorously controlled for, apparent effects of trait-based dehumanization are better explained by a relative preference for the ingroup compared to criminals. Criminals were rated as 'less human' on the ascent of man scale, showing we should expect to see trait-based dehumanization should the process occur.

Taken together, Study 3 calls into question prior work that reports dehumanization of criminals within the dual model framework (Bastian et al., 2013; Viki et al., 2012). These studies purportedly show that dehumanization of criminals predicts outcomes such as sentencing judgements and support for rehabilitation programmes. However, our results suggest these studies may not have accurately measured dehumanization. Rather, because of a desirability bias within the human specific attributions made, these studies are likely to have inadvertently measured dislike and negative feeling towards the offenders in question rather than dehumanization. Thus, it may be more accurate to conclude from these studies that negative feeling towards (rather than trait-based dehumanization of) certain offenders predicts judicial outcomes. Results from Study 3 are important because accurately understanding the mechanisms underlying biases in character judgements towards offenders has implications for the success of criminal justice programmes.

6. Analysis of combined data

Taken together, the reported studies suggest that there is no support for the dual model of dehumanization in these intergroup contexts. However, a remaining possibility is that trait-based dehumanization does occur in these contexts, albeit with a smaller effect size than we were powered to detect. This seems unlikely as previous research on trait-based dehumanization has reported large effect sizes with fewer participants (e.g, Bain et al., 2009; Bastian et al., 2013; Loughnan & Haslam, 2007) and we pre-registered our sample size to detect medium

effect sizes. Nevertheless, to investigate this possibility further, we conducted two exploratory analyses to determine whether there was any evidence for animalistic or mechanistic dehumanization when the data from studies 2 and 3 (which had an identical design) were combined. This resulted in two additional analyses each with a sample size of 260 and power to detect interactions with a much smaller η_p^2 of 0.03.

A 2 (target group: ingroup/outgroup) x 2 (humanness: uniquely human/shared with other species) x 2 (valence: desirable/undesirable) x 2(Experiment: 2a/3a) mixed ANOVA tested for interactions showing animalistic dehumanization as separable from ingroup preference. There were significant two-way interactions between group and valence, $F(1, 258) = 124.17$, $p < .001$, $\eta_p^2 = 0.325$ and also between group and humanness, $F(1, 258) = 10.68$, $p = .001$, $\eta_p^2 = 0.040$. Note that this 2 way interaction did not reflect the predictions of the dual model - ratings were higher for ingroup than outgroup on both uniquely human traits ($p < .001$) and those shared with other species ($p = .028$). Importantly, these effects were qualified in a three-way interaction between group, humanness and valence, $F(1, 258) = 8.56$, $p = .004$, $\eta_p^2 = 0.032$. This showed that ratings were significantly greater for ingroup than outgroup on desirable traits, both uniquely human and shared with other species, but significantly greater for outgroup than ingroup on undesirable traits, both uniquely human and shared with other species (all $ps < 0.001$). The four-way interaction between group, humanness, valence and experiment was not significant, $F(1, 258) = 2.90$, $p = .090$, $\eta_p^2 = 0.011$.

A further 2 (target group: ingroup/outgroup) x 2 (humanness: uniquely human/shared with robots) x 2 (valence: desirable/undesirable) x 2(Experiment: 2b/3b) mixed ANOVA tested for interactions showing mechanistic dehumanization as separable from ingroup preference. There were significant two-way interactions between group and valence, $F(1, 258) = 190.26$, $p < .001$, $\eta_p^2 = 0.424$ and also between group and humanness, $F(1, 258) = 8.14$, $p = .005$, $\eta_p^2 = 0.031$. Note that once again this interaction did not reflect the predictions of the dual model - ratings were higher for ingroup than outgroup on both uniquely human traits and those shared with robots ($p < .001$). Importantly, these were qualified in a three-way interaction between group, humanness and valence, $F(1, 258) = 108.49$, $p < .001$, $\eta_p^2 = 0.296$. These showed that ratings were significantly greater for ingroup than outgroup on desirable traits, both uniquely human and shared with robots ($ps < 0.001$), but significantly greater for outgroup than ingroup on undesirable traits, both uniquely human ($p < .001$) and shared with robots ($p = .004$). The four-way interaction between group, humanness, valence and experiment was this time significant, $F(1, 258) = 14.00$, $p < .001$, $\eta_p^2 = 0.051$. This interaction reflected the finding that in Experiment 2b (immigrant outgroup) attributions of undesirable traits shared with robots was only marginally greater for outgroup than ingroup ($p = .068$), whilst in Experiment 3b (criminal outgroup) this effect was stronger ($p = .022$).

7. General discussion

In this paper, we question the central claims of one of the most prominent psychological accounts of dehumanization - the dual model - which holds that outgroup members are perceived as lesser humans than ingroup members by being denied human specific traits (Haslam, 2006). We first revisited work relating to how the lay concept of 'human' is best characterised. We then tested its predictions about outgroup dehumanization in a series of seven experiments. Our results present a serious empirical challenge to the dual model.

The dual model argues that there are two sense of humanness: human uniqueness and human nature. Uniquely human traits can be summarised as civility, refinement, moral sensibility, rationality, and maturity. Human nature traits can be summarised as emotional responsiveness, interpersonal warmth, cognitive openness, agency, and depth (Haslam, 2006). However, the traits that supposedly characterise 'humanness' within this model are broadly socially desirable (Over, 2020a; Over, 2020b). We showed that people also associate some undesirable traits with the concept 'human'. As well as considering humans to be refined and cultured, people also consider humans to be corrupt, selfish and cruel.

Results from our pretest provided us with grounds for re-examining predictions made by the dual model of dehumanization about the nature of intergroup bias in trait attributions. The dual model account holds that lesser attribution of human specific traits to outgroup members represents a psychological process of dehumanization that is separable from ingroup preference. However, as the human specific attributes summarised by the model are positive and socially desirable, it is possible that previous findings are better explained in terms of ingroup preference, the process of attributing positive qualities to ingroup members to a greater extent than to outgroup members.

In seven highly-powered experiments, we tested the predictions of the dual model against this alternative. We pitted the two hypotheses against each other by comparing attributions of uniquely human traits that varied in whether they were socially desirable or undesirable to ingroup and outgroup members. The dual model holds that subtle dehumanization is evidenced by denying outgroup members uniquely human traits relative to ingroup members. We reasoned that whereas outgroup members may be denied desirable human traits, they are likely to be attributed undesirable human traits to a greater extent than ingroup members.

Across three distinct intergroup contexts, we found no evidence for either animalistic or mechanistic dehumanization of outgroup members. Instead, we found strong and reliable intergroup preference effects. Desirable traits were ascribed more strongly to ingroup members than outgroup members and undesirable traits more strongly to outgroup members than ingroup members, irrespective of perceived humanness.

A possible defence of the dual model account could be to argue that we chose three intergroup contexts in which animalistic and mechanistic dehumanization does not occur. However, we chose to investigate judgements of political opponents, immigrants and criminals specifically because previous research has suggested that they are dehumanized on a range of measures (Banton et al., 2020; Bastian et al., 2013; Markowitz & Slovic, 2020; Pacilli et al., 2016; Viki et al., 2013). In addition, we also showed in every experiment that outgroup members were explicitly rated as less human than were the ingroup on the blatant dehumanization scale (Kteily et al., 2015). Prior work shows that measures of animalistic and mechanistic dehumanization correlate positively with blatant dehumanization scores (Kteily et al., 2015). Though they are not claimed to measure the same construct, they have been shown to reliably co-occur. These findings confirm that these are the sorts of intergroup contexts in which we would expect to see trait-based dehumanization should the process occur.

We acknowledge that without testing all possible intergroup contexts, it remains a possibility that some outgroups could be denied human specific attributes relative to ingroups even when valence is appropriately controlled for. In other words, it could be the case that trait-based dehumanization occurs independently of ingroup preference in some social settings. It may be particularly interesting for future research to investigate intergroup contexts that are not so strongly associated with competition, threat and animosity.

However, the possibility that some, as yet untested, groups may be denied human unique attributes does not detract from the importance of our critique. To accurately measure trait-based dehumanization in future research, studies must consider the central role of valence. Prior work utilising the dual model framework has reported dehumanization to be extremely widespread in society, affecting not just marginalised groups but doctors, patients and even cyclists (Delbosc, Naznin, Haslam, & Haworth, 2019; Haslam & Stratemeyer, 2016). Rigorous measurement and tighter experimental control may change some or all of the conclusions from previous research.

Across our experiments, we observed strong intergroup preference effects, with desirable traits more strongly ascribed to the ingroup and

undesirable traits more strongly to the outgroup. Our results demonstrate both ingroup favouritism (assigning greater positivity to the ingroup) and outgroup derogation (assigning greater negativity to the outgroup) (Brewer, 1999; Hewstone et al., 2002). However, we also suggest that group specific stereotypes are likely to play an important role in these processes. In many social contexts, trait attributions may reflect social stereotyping as well as intergroup preferences (Fiske et al., 2002). For example, previous work suggested that Anglo-Australians were 'animalistically' dehumanized both by themselves and by Ethnic-Chinese participants, whilst Ethnic-Chinese people were 'mechanistically' dehumanized both by themselves and by Anglo-Australians (Bain et al., 2009). These effects may be more compatible with stereotype content than with trait-based dehumanization. Future work would benefit from addressing the distinction between stereotyping and trait-based dehumanization.

An outstanding question relates to whether other psychological models of dehumanization more accurately capture the ways in which different social groups are perceived. For example, infrahumanisation theory predicts that people tend to believe ingroup members experience uniquely human emotions more strongly than do outgroup members (Leyens et al., 2000, 2001). It would be valuable for future research to examine the utility of this theory by testing whether participants perceive ingroup members to experience human emotions more strongly overall or whether they perceive ingroup members to experience pro-social emotions more strongly but outgroup members to experience antisocial emotions more strongly. Further work could helpfully investigate how these findings bare on the claim that outgroups are sometimes dehumanized by being denied mental states (Harris & Fiske, 2006).

Taken together, our studies suggest that the dual model does not accurately characterise the ways in which outgroups are perceived in at least the social contexts examined here – political groups, immigrants and criminals. Prejudice and discrimination are pressing social problems. If psychological research is to contribute to the interdisciplinary mission to reduce prejudice and encourage more egalitarian behaviour, then it must start by accurately characterising the psychological biases underlying discriminatory behaviour. We suggest that the dual model of dehumanization conflates apparent evidence for dehumanization with ingroup preference. As a result, it may obscure more than it reveals about the psychology of intergroup bias.

Credit author statement

FE, HO and ST conceptualised and designed the experiments. FE collected the data and FE and JF performed the analyses. FE and HO wrote the original draft. All authors edited and reviewed the manuscript.

Declaration of Competing Interest

None.

Acknowledgements

This research was supported by the European Research Council under the European Union's Horizon 2020 Programme, grant number ERC-STG- 755719, awarded to HO. We would like to thank David Livingstone Smith and two anonymous reviewers for valuable comments on an earlier draft.

References

Andrighetto, L., Baldissarri, C., Lattanzio, S., Loughnan, S., & Volpato, C. (2014). Human-itarian aid? Two forms of dehumanization and willingness to help after natural disasters. *British Journal of Social Psychology, 53*(3), 573–584.

Bain, P., Park, J., Kwok, C., & Haslam, N. (2009). Attributing human uniqueness and human nature to cultural groups: Distinct forms of subtle dehumanization. *Group Processes & Intergroup Relations, 12*(6), 789–805. https://doi.org/10.1177/1368430209340415.

Banton, O., West, K., & Kinney, E. (2020). The surprising politics of anti-immigrant prejudice: How political conservatism moderates the effect of immigrant race and religion on infrahumanization judgements. *British Journal of Social Psychology, 59*(1), 157–170. https://doi.org/10.1111/bjso.12337.

Bastian, B., Denson, T. F., & Haslam, N. (2013). The roles of dehumanization and moral outrage in retributive justice. *PLoS One, 8*(4).

Bloom, P. (2017 November 20). The root of all cruelty? *The New Yorker*. Retrieved from https://www.newyorker.com/magazine/2017/11/27/the-root-of-all-cruelty.

Brewer, M. B. (1999). The psychology of prejudice: Ingroup love and outgroup hate? *Journal of Social Issues, 55*(3), 429–444.

Bruneau, E., Kteily, N., & Laustsen, L. (2018). The unique effects of blatant dehumanization on attitudes and behavior towards Muslim refugees during the European 'refugee crisis' across four countries. *European Journal of Social Psychology, 48*(5), 645–662. https://doi.org/10.1002/ejsp.2357.

Cassese, E. C. (2020). Dehumanization of the opposition in political campaigns. *Social Science Quarterly, 101*(1), 107–120.

Delbosc, A., Naznin, F., Haslam, N., & Haworth, N. (2019). Dehumanization of cyclists predicts self-reported aggressive behaviour toward them: A pilot study. *Transportation Research Part F: Traffic Psychology and Behaviour, 62*, 681–689.

Demoulin, S., Leyens, J., Paladino, M., Rodriguez-Torres, R., Rodriguez-Perez, A., & Dovidio, J. (2004). Dimensions of "uniquely" and "non-uniquely" human emotions. *Cognition & Emotion, 18*(1), 71–96. https://doi.org/10.1080/02699930244000444.

Demoulin, S., Pozo, B. C., & Leyens, J.-P. (2009). Infrahumanization: The differential interpretation of primary and secondary emotions. In S. Demoulin, J.-P. Leyens, & J. F. Dovidio (Eds.), *Vol. 1. Intergroup misunderstandings: Impact of divergent social realities* (pp. 153–172). Psychology Press.

Fasoli, F., Paladino, M. P., Carnaghi, A., Jetten, J., Bastian, B., & Bain, P. G. (2016). Not "just words": Exposure to homophobic epithets leads to dehumanizing and physical distancing from gay men. *European Journal of Social Psychology, 46*(2), 237–248.

Fernández-Reino, M. (2019). *The health of migrants in the UK*.

Fernández-Reino, M. (2020). *Migrants and discrimination in the UK*. The Migrant Observatory at the University of Oxford.

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

Goodwin, M., & Milazzo, C. (2017). Taking back control? Investigating the role of immigration in the 2016 vote for Brexit. *The British Journal of Politics and International Relations, 19*(3), 450–464.

Harris, L., & Fiske, S. (2011). Dehumanized perception a psychological means to facilitate atrocities, torture, and genocide? *Zeitschrift Für Psychologie/Journal of Psychology, 219*, 175–181. https://doi.org/10.1027/2151-2604/a000065.

Harris, L. T., & Fiske, S. T. (2006). Dehumanizing the lowest of the low: Neuroimaging responses to extreme out-groups. *Psychological Science, 17*(10), 847–853. https://doi.org/10.1111/j.1467-9280.2006.01793.x.

Haslam, N. (2006). Dehumanization: An integrative review. *Personality and Social Psychology Review, 10*(3), 252–264. https://doi.org/10.1207/s15327957pspr1003_4.

Haslam, N. (2019). *The Many Roles of Dehumanization in Genocide* (pp. 119–138). https://doi.org/10.1093/oso/9780190685942.003.0005.

Haslam, N., Bain, P., Douge, L., Lee, M., & Bastian, B. (2005). More human than you: Attributing humanness to self and others. *Journal of Personality and Social Psychology, 89*(6), 937–950. https://doi.org/10.1037/0022-3514.89.6.937.

Haslam, N., Bastian, B., & Bissett, M. (2004). *Essentialist beliefs about personality and their implications*. https://doi.org/10.1177/0146167204271182.

Haslam, N., Kashima, Y., Loughnan, S., Shi, J., & Suitner, C. (2008). Subhuman, inhuman, and superhuman: Contrasting humans with nonhumans in three cultures. *Social Cognition, 26*(2), 248–258. https://doi.org/10.1521/soco.2008.26.2.248.

Haslam, N., & Loughnan, S. (2014). Dehumanization and Infrahumanization. *Annual Review of Psychology, 65*(1), 399–423. https://doi.org/10.1146/annurev-psych-010213-115045.

Haslam, N., & Loughnan, S. (2016). How dehumanization promotes harm. *The Social Psychology of Good and Evil, 2*, 140–158.

Haslam, N., & Stratemeyer, M. (2016). Recent research on dehumanization. *Current Opinion in Psychology, 11*, 25–29. https://doi.org/10.1016/j.copsyc.2016.03.009.

Hewstone, M., Rubin, M., & Willis, H. (2002). Intergroup bias. *Annual Review of Psychology, 53*, 575–604. https://doi.org/10.1146/annurev.psych.53.100901.135109.

Hodson, G., & Costello, K. (2007). Interpersonal disgust, ideological orientations, and dehumanization as predictors of intergroup attitudes. *Psychological Science, 18*(8), 691–698.

Kauff, M., Wölfer, R., & Hewstone, M. (2017). Impact of discrimination on health among adolescent immigrant minorities in Europe: The role of perceived discrimination by police and security personnel. *Journal of Social Issues, 73*(4), 831–851.

Kteily, N., & Bruneau, E. (2017). Backlash: The politics and real-world consequences of minority group dehumanization. *Personality and Social Psychology Bulletin, 43*(1), 87–104.

Kteily, N., Bruneau, E., Waytz, A., & Cotterill, S. (2015). The ascent of man: Theoretical and empirical evidence for blatant dehumanization. *Journal of Personality and Social Psychology, 109*(5), 901–931. https://doi.org/10.1037/pspp0000048.

Lang, J. (2010). Questioning dehumanization: Intersubjective dimensions of violence in the Nazi concentration and death camps. *Holocaust and Genocide Studies, 24*, 225–246. https://doi.org/10.1093/hgs/dcq026.

Lang, J. (2020). The limited importance of dehumanization in collective violence. *Current Opinion in Psychology, 35*, 17–20.

Leidner, B., Castano, E., & Ginges, J. (2013). Dehumanization, retributive and restorative justice, and aggressive versus diplomatic intergroup conflict resolution strategies. *Personality and Social Psychology Bulletin, 39*(2), 181–192.

Leyens, J., Demoulin, S., Vaes, J., Gaunt, R., & Paladino, M. P. (2007). Infra-humanization: The wall of group differences. *Social Issues and Policy Review, 1*(1), 139–172. https://doi.org/10.1111/j.1751-2409.2007.00006.x.

Leyens, J.-P. (2009). Retrospective and prospective thoughts about infrahumanization. *Group Processes & Intergroup Relations, 12*(6), 807–817.

Leyens, J.-P., Cortes, B., Demoulin, S., Dovidio, J. F., Fiske, S. T., Gaunt, R., … Vaes, J. (2003). Emotional prejudice, essentialism, and nationalism the 2002 Tajfel lecture. *European Journal of Social Psychology, 33*(6), 703–717. https://doi.org/10.1002/ejsp.170.

Leyens, J.-P., Paladino, P. M., Rodriguez-Torres, R., Vaes, J., Demoulin, S., Rodriguez-Perez, A., & Gaunt, R. (2000). The emotional side of prejudice: The attribution of secondary emotions to Ingroups and Outgroups. *Personality and Social Psychology Review, 4*(2), 186–197. https://doi.org/10.1207/S15327957PSPR0402_06.

Leyens, J.-P., Rodriguez-Perez, A., Rodriguez-Torres, R., Gaunt, R., Paladino, M.-P., Vaes, J., & Demoulin, S. (2001). Psychological essentialism and the differential attribution of uniquely human emotions to ingroups and outgroups. *European Journal of Social Psychology, 31*(4), 395–411. https://doi.org/10.1002/ejsp.50.

Loughnan, S., & Haslam, N. (2007). Animals and androids: Implicit associations between social categories and nonhumans. *Psychological Science, 18*(2), 116–121. https://doi.org/10.1111/j.1467-9280.2007.01858.x.

Macrae, C. N., & Hewstone, M. R. C. (1990). Cognitive biases in social categorization: Process and consequences. In J.-P. Caverni, J.-M. Fabre, & M. Gonzalez (Eds.), *vol. 68*. *Advances in psychology* (pp. 325–348). North-Holland. https://doi.org/10.1016/S0166-4115(08)61331-X.

Manne, K. (2016). Humanism: A critique. *Social Theory and Practice, 42*(2), 389–415 (JSTOR).

Manne, K. (2018). *Down girl: The logic of misogyny*. Oxford University Press.

Markowitz, D. M., & Slovic, P. (2020). Social, psychological, and demographic characteristics of dehumanization toward immigrants. *Proceedings of the National Academy of Sciences, 117*(17), 9260–9269.

Medin, D. L., & Smith, E. E. (1984). Concepts and concept formation. *Annual Review of Psychology, 35*, 113–138. https://doi.org/10.1146/annurev.ps.35.020184.000553.

Meleady, R., Seger, C. R., & Vermue, M. (2017). Examining the role of positive and negative intergroup contact and anti-immigrant prejudice in B rexit. *British Journal of Social Psychology, 56*(4), 799–808.

Over, H. (2020a). Seven challenges for the dehumanization hypothesis. *Perspectives on Psychological Science, 16*(1), 3–13. https://doi.org/10.1177/1745691620902133.

Over, H. (2020b). Falsifying the dehumanization hypothesis. *Perspectives on Psychological Science, 16*(1), 33–38. https://doi.org/10.1177/1745691620969657.

Pacilli, M. G., Roccato, M., Pagliaro, S., & Russo, S. (2016). From political opponents to enemies? The role of perceived moral distance in the animalistic dehumanization of the political outgroup. *Group Processes & Intergroup Relations, 19*(3), 360–373.

Paladino, M.-P., Leyens, J.-P., Rodriguez, R., Rodriguez, A., Gaunt, R., & Demoulin, S. (2002). Differential Association of Uniquely and non Uniquely Human Emotions with the Ingroup and the Outgroup. *Group Processes & Intergroup Relations, 5*(2), 105–117. https://doi.org/10.1177/1368430202005002539.

Phillips, J., & Cushman, F. (2017). Morality constrains the default representation of what is possible. *Proceedings of the National Academy of Sciences, 114*(18), 4649–4654.

Smith, D. L. (2011). *Less than human: Why we demean, enslave, and exterminate others*. St. Martin's Publishing Group.

Smith, D. L. (2014). Dehumanization, essentialism, and moral psychology: Dehumanization, essentialism, and moral psychology. *Philosophy Compass, 9*(11), 814–824. https://doi.org/10.1111/phc3.12174.

Smith, D. L. (2016). Paradoxes of dehumanization. *Social Theory and Practice, 42*(2), 416–443. https://doi.org/10.5840/soctheorpract201642222.

Smith, E. E., & Medin, D. L. (1981). *Categories and concepts* (Vol. 9). MA: Harvard University Press Cambridge.

Tirrell, L. (2012). Genocidal language games. In I. Maitra, & M. K. McGowan (Eds.), *Speech and Harm: Controversies Over Free Speech (pp. 174–221)*. Oxford University Press.

Vaes, J., Leyens, J.-P., Paola Paladino, M., & Pires Miranda, M. (2012). We are human, they are not: Driving forces behind outgroup dehumanization and the humanisation of the ingroup. *European Review of Social Psychology, 23*(1), 64–106. https://doi.org/10.1080/10463283.2012.665250.

Vaes, J., Paladino, M. P., Castelli, L., Leyens, J.-P., & Giovanazzi, A. (2003). On the behavioral consequences of Infrahumanization: The implicit role of uniquely human emotions in intergroup relations. *Journal of Personality and Social Psychology, 85*(6), 1016–1034. https://doi.org/10.1037/0022-3514.85.6.1016.

Vaes, J., Paladino, M.-P., & Leyens, J.-P. (2002). The lost e-mail: Prosocial reactions induced by uniquely human emotions. *British Journal of Social Psychology, 41*(4), 521–534. https://doi.org/10.1348/014466602321149867.

Viki, G. T., Fullerton, I., Raggett, H., Tait, F., & Wiltshire, S. (2012). The role of dehumanization in attitudes toward the social exclusion and rehabilitation of sex offenders. *Journal of Applied Social Psychology, 42*(10), 2349–2367.

Viki, G. T., Osgood, D., & Phillips, S. (2013). Dehumanization and self-reported proclivity to torture prisoners of war. *Journal of Experimental Social Psychology, 49*(3), 325–328. https://doi.org/10.1016/j.jesp.2012.11.006.

Viki, G. T., Winchester, L., Titshall, L., Chisango, T., Pina, A., & Russell, R. (2006). Beyond secondary emotions: The Infrahumanization of Outgroups using human–related and animal–related words. *Social Cognition, 24*(6), 753–775. https://doi.org/10.1521/soco.2006.24.6.753.

Yee, E., & Thompson-Schill, S. L. (2016). Putting concepts into context. *Psychonomic Bulletin & Review, 23*(4), 1015–1027. https://doi.org/10.3758/s13423-015-0948-7.