Main manuscript for:

Intergroup preference, not dehumanization, explains social biases in emotion attribution

Authors: Florence E. Enock[1], Steven P. Tipper[1], and Harriet Over[1]

[1]Department of Psychology, University of York, York, United Kingdom. YO10 5DD.

*Correspondence to: Florence Enock, Florence.Enock@york.ac.uk.

Author contact details

Florence Enock: Florence.enock@York.ac.uk, 01904 322902

Steven Tipper: Steven.Tipper@York.ac.uk, 01904 323162

Harriet Over: Harriet.Over@York.ac.uk, 01904 322906

Acknowledgments

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.

Author Contributions

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

Competing interests

The authors declare no competing interests.

Abstract

Psychological models can only help improve intergroup relations if they accurately characterise the mechanisms underlying social biases. The claim that outgroups suffer dehumanization is near ubiquitous in the social sciences. We challenge the most prominent psychological model of dehumanization - infrahumanization theory - which holds outgroup members are subtly dehumanized by being denied human emotions. We examine the theory across seven intergroup contexts in thirteen pre-registered and highly powered experiments (*N*=1,690). We find outgroup members are not denied uniquely human emotions relative to ingroup members. Rather, they are ascribed prosocial emotions to a lesser extent but antisocial emotions to a greater extent. Apparent evidence for infrahumanization is better explained by ingroup preference, outgroup derogation and stereotyping. Infrahumanization theory may obscure more than it reveals about intergroup bias.

Keywords: Dehumanization, infrahumanization, social cognition, intergroup bias, prejudice.

1. Introduction

The claim that outgroup members are perceived as 'less than human' has been extremely influential in social psychology, social neuroscience, philosophy and sociology. It has entered into public rhetoric as well, regularly being discussed in the media. Blatant forms of dehumanization are thought to reveal themselves in propaganda and other forms of hate speech in which outgroup members are described as less than human creatures, for example as similar to rats, parasites or vermin (Haslam, 2006; Smith, 2011). Blatant dehumanization has been linked to extreme intergroup harm such as genocide, torture and police brutality towards African Americans (Goff et al., 2008, 2014; Smith, 2011; Tirrell, 2012).

Subtler forms of dehumanization, in which outgroups are considered somewhat less human, are hypothesised to be widespread and are typically studied in lab-based settings (Harris & Fiske, 2006; Haslam, 2006; Leyens et al., 2000, 2001). In subtle forms of dehumanization, outgroup members are thought to possess uniquely human qualities to a lesser extent than do the ingroup. Three psychological models of subtle dehumanization have been particularly prominent. According to Harris & Fiske (2006), to the extent outgroups are dehumanized, they are thought to possess mental states to a lesser extent than do the ingroup. According to the dual model, outgroups are thought to possess uniquely human character traits to a lesser extent than do the ingroup (Haslam, 2006). According to infrahumanization theory, perhaps the most prominent of the three models, outgroups are thought to possess uniquely human emotions to a lesser extent than do the ingroup (Leyens et al., 2000, 2001). These subtle forms of dehumanization have been linked to negative outcomes including reduced prosocial behaviour

towards outgroups (Cuddy et al., 2007; Vaes et al., 2003). In a world of social division, with frequently occurring cases of discrimination based on religion, ethnicity and gender, to name only a few, understanding the extent and consequences of dehumanization is crucial.

In recent years, several theoretical critiques of research on dehumanization have emerged (Appiah, 2008; Bloom, 2017; Lang, 2010, 2020; Manne, 2016, 2018; Over, 2020a, 2020b; Rai et al., 2017; Smith, 2014, 2020). These critiques suggest that perceiving outgroups as 'less than human' might be less common than it first appears. Considering blatant dehumanization, as evidenced by historical examples of propaganda, Manne (2016) and Bloom (2017) have both pointed out that victims said to be 'dehumanized' are often described with terms that only really make sense when applied to humans, albeit negative and antisocial ones. For example, in Nazi propaganda, Jewish people were frequently described as ruthless, corrupt, treacherous and criminally minded, terms out of place when used to describe an animal or a machine. Relatedly, the hypothesised causal connection between dehumanization and intergroup harm has been questioned. Several theorists suggest that being perceived as having certain human qualities, such as being corrupt, spiteful or deceptive, may actually increase people's risk of harm (Appiah, 2008; Bloom, 2017; Lang, 2010, 2020; Manne, 2016, 2018; Over, 2020a, 2020b). Empirical research showing that morally-motivated intergroup harm is not linked to dehumanization lends support to these critiques (Rai et al., 2017).

Turning to lab-based research, Over (2020a, 2020b) argues that what appears to be evidence for dehumanization, as operationalised by the dual model (Haslam, 2006) and infrahumanization theory (Leyens et al., 2000; 2001), may be

49 better explained in terms of intergroup preference effects (a general tendency to
50 prefer the ingroup to the outgroup). According to the dual model of dehumanization,
51 outgroup members tend to be attributed uniquely human character traits to a lesser
52 extent than are the ingroup (Haslam, 2006). However, to date, the overwhelming
53 majority of the traits included in empirical research are socially desirable, for
54 example, warmth, rationality, civility and refinement. Over (2020a; 2020b)
55 hypothesises that while outgroup members may be thought to possess some
56 uniquely human qualities to a lesser extent, for example, civility, refinement and
57 rationality, there may be other uniquely human qualities that are more strongly
58 attributed to the outgroup than the ingroup. For example, antisocial human
59 characteristics such as jealousy, arrogance and bitterness seem to only make sense
60 in the context of humans but are unlikely to be attributed more strongly to ingroup
61 than outgroup members.
62 Recent experimental work from Enock and colleagues (2021) supports
63 Over's critique, presenting an empirical challenge to the dual model of
64 dehumanization (Haslam, 2006). Enock et al. (2021) first established that people
65 tend to associate undesirable characteristics as well as desirable ones with humans,
66 confirming an omission from the dual model. Subsequently, seven experiments
67 tested the predictions of the dual model directly against a social preference account
68 in three distinct intergroup contexts - political opponents, immigrants and criminals.
69 Results showed no evidence for dehumanization when undesirable as well as
70 desirable human traits were included in the stimuli. Rather, in line with the social
71 preference account, desirable traits were ascribed more strongly to ingroup
72 members than outgroup members and undesirable traits more strongly to outgroup

73 members than ingroup members, irrespective of perceived humanness (Enock et
74 al., 2021).
75 Perhaps the most prominent psychological model of dehumanization is
76 infrahumanization theory (Leyens et al., 2000, 2001). This theory is distinct from the
77 dual model in that it proposes outgroup members are denied uniquely human
78 emotions rather than character traits. The model is founded on the widespread
79 notion that there is a distinction between secondary emotions (such as pride and
80 guilt) and primary emotions (such as happiness and anger). The former are
81 considered unique to humans, the latter shared with other animals (Demoulin et al.,
82 2004; Ekman, 1992; Leyens et al., 2000). Seminal work has found that when
83 choosing emotions to best describe different groups, people preferentially ascribe
84 uniquely human emotions more strongly to ingroup members (Leyens et al., 2001).
85 For example, across a range of social contexts, participants ascribed uniquely
86 human emotions such as hope, compassion, pride, melancholy, disappointment and
87 remorse, to ingroup members to a greater extent than to outgroup members (Banton
88 et al., 2020; Cortes et al., 2005; Leyens et al., 2001; Paladino et al., 2002; Prati et
89 al., 2016)
90 Infrahumanization research has proliferated in recent years (Leyens, 2009;
91 Vaes et al., 2012). Effects have been reported across explicit and implicit measures
92 (Boccato et al., 2007; Paladino et al., 2002), and a multitude of intergroup contexts,
93 including regional, religious and racial identities (Banton et al., 2020; Rodríguez-
94 Pérez et al., 2011), university affiliations (Vaes et al., 2003) and minimal groups
95 (Demoulin et al., 2009; Simon & Gutsell, 2020). The importance of the model is
96 highlighted by the use of infrahumanization as an outcome measure in interventions

to improve intergroup relations (Brown et al., 2007; Capozza et al., 2013; Prati et al., 2016; Tam et al., 2007).

Of key importance to infrahumanization theory is the claim that subtle dehumanization is distinct from intergroup preference because participants ascribe both positive (e.g., hope, admiration) and negative (e.g., guilt, remorse) uniquely human emotions more strongly to the ingroup than the outgroup (Leyens et al., 2000, 2001). That the effect is observed for negative human emotions is claimed to be crucial for separating infrahumanization from intergroup preference. According to infrahumanization theory, the process of perceiving others as lacking *negative* human emotions shows a subtle form of dehumanization that is separate from derogation because here, 'humanness' is different from 'good' (Castano & Giner-Sorolla, 2006; Haslam & Loughnan, 2014; Haslam & Stratemeyer, 2016). As Leyens and colleagues (2000, p.189) note in their original hypotheses:

> *People should more easily associate their ingroup than an outgroup with secondary emotions. This preferential association should be true independent of the valence of the secondary emotions. Indeed, it is the category of secondary emotions as such that is considered typically human. No qualification is made for positive or negative secondary emotions.*

Explaining this further, Leyens and colleagues (2001, p. 398) argue:

> *If the attribution of secondary emotions to the ingroup reflected a mere positivity effect, it would lose its interest and originality… would people also select more negative secondary emotions for their ingroup than for the outgroup? A positive answer… would get rid of a simple positivity bias in the case of secondary emotions.*

In their original studies, Leyens et al. (2001) included valence (positive/negative) as a factor in their analysis. They found that more secondary emotions were attributed to the ingroup than the outgroup and that this effect was not qualified by the valence of the emotions (see Leyens et al., 2001, p. 402, Fig. 2). Subsequent work from Cortes and colleagues (2005) demonstrated similar effects, showing more positive (e.g., contentment, delight) and negative (e.g., melancholy, resignation) uniquely human emotions to be ascribed to the ingroup than the outgroup, again not qualified by emotion valence (see Cortes et al., 2005, p. 247, Figure 1). Convergent evidence is provided by research that treats and valence and humanness as continuous factors (Castano & Gina-Sorolla, 2006) and by research that has measured the attribution of positive (e.g., amazement, compassion) and negative (e.g., despair, guilt) emotions in separate experiments (Paladino et al., 2002).

Owing to the inclusion of negative as well as positive emotions in infrahumanization research, the theory thus appears to be immune to Over's (2020a; 2020b) critique. However, we reconsider infrahumanization theory through a framework that understands emotions as social as well as individual experiences (Parkinson, 1996). Some emotions are by definition positive to experience but are somewhat unkind to others (e.g., schadenfreude), while some may be negative to experience but are not inherently unkind to others (e.g., disappointment). We suggest that although emotions such as guilt and remorse are negative to experience, they are not necessarily antisocial in character. Rather, they indicate appropriate responses to moral wrong-doing and thus people who display them are viewed positively (Stearns & Parrott, 2012). Understanding emotion as a social

phenomenon raises the conceptual distinction between valence (positive or negative to experience) and sociality (prosocial or antisocial as viewed by others).

To our knowledge, no work has yet measured whether previously reported infrahumanization effects are independent of emotion sociality. Terms frequently included as exemplars of negative uniquely human emotions in infrahumanization work such as melancholy, guilt and remorse, may be negative to experience, but are not obviously antisocial. This omission makes it impossible to determine whether infrahumanization really is separable from ingroup preference and thus whether it holds unique explanatory value in intergroup relations.

Twenty years since infrahumanization theory was proposed, we revisit and test its founding claims in thirteen pre-registered, highly powered experiments. In our first six experiments (Study 1) we show that previously reported infrahumanization effects broadly replicate across multiple intergroup contexts. In six subsequent experiments (Study 2), we remove the confound in previous research by introducing emotions that differ in sociality rather than valence. In line with the social preference account, we show that apparent evidence for infrahumanization is better explained by ingroup preference and stereotyping. In a final experiment (Study 3), we provide further evidence for the social preference account by replicating the pattern of results observed in Study 2 in a minimal group design.

In Studies 1 and 2, we use the same six intergroup contexts. The precise social conditions necessary for infrahumanization have not been clearly established within the field and it has been noted that it may not always occur (e.g., Castano & Giner-Sorolla, 2006; Demoulin et al., 2009). However, a comprehensive review of

prior empirical evidence suggests that outgroups are particularly likely to be infrahumanized if they threaten one's worldview, are disliked, and belong to a social category that one would not want to belong to (Leyens, 2009). Initial infrahumanization studies included students from the Canary Islands versus those from mainland Spain as the intergroup context (Leyens et al., 2000, 2001). The researchers noted general hostility between these groups, with each seeing the other as a 'disliked' outgroup, suggesting 'outgroup derogation' was likely (Leyens et al., 2001, pp. 396–399). Follow-up work included Spanish or Belgian ingroup members and North African individuals as outgroup members, at the time a 'very stigmatized minority and low-status group in Belgium and in the Canary Islands' (Paladino et al., 2002, p. 113). Most infrahumanization studies focus on social contexts that are similarly grounded in antagonism (e.g., Banton et al., 2020; Gaunt, 2009).

We chose our groups to maximise our chances of replicating infrahumanization effects if they occur (Leyens, 2009). The first outgroup we chose was Muslims (Christian ingroup) (Expts. 1a&2a). Dehumanization of religious outgroups, including of Muslims by Christians, has been widely reported (Banton et al., 2020; Kteily et al., 2016; Viki et al., 2013) and discrimination against Muslims is a pressing social problem in many Western societies (Calfano, 2018; Hewstone & Schmid, 2014). The remaining outgroups were criminals (Expts. 1b&2b), child molesters (Expts. 1c&2c), anti-vaxxers (Expts. 1d&2d), people who do not adhere to social distancing regulations during the Covid-19 pandemic ('non-social distancers') (Expts. 1e&2e), and climate change deniers (Expts. 1f&2f). Prior work reports dehumanization of criminals and sex offenders (Bastian et al., 2013; Viki et al.,

2012). We introduced the three additional outgroup contexts (anti-vaxxers, non-social distancers and climate change deniers) based on current pertinence. At the time of data collection in early April 2020, the UK had been in full 'lockdown' for just over one week and tension between individuals who did and did not adhere to the guidelines was developing (Prosser et al., 2020). Similarly, social division between those who are pro- and anti-vaccination has been particularly salient during the COVID-19 pandemic (Johnson et al., 2020). Rather than seeking to be exhaustive, the intergroup contexts we chose for Studies 1 and 2 illustrate the conceptual distinction between infrahumanization theory and our alternative social preference account. In Study 3, we replicated our results in a minimal group design. This allowed us to further demonstrate the generalisability of our results in a social context free from prior stereotypes and intergroup antagonism.

2. Data collection and availability

All experiments reported in this manuscript took place online and were created and administered using Qualtrics (https://www.qualtrics.com). Participants were recruited through Prolific (https://www.prolific.co) and a different sample was included for each experiment reported. Informed consent was obtained at the start of each session according to approved ethical procedures. Participants were compensated at an approximate rate of £7.50 per hour. All studies were pre-registered and the data is available open access. Links to pre-registration documents and raw data for each study can be found at: https://osf.io/rzb3n/

3. Pretest

One concern with prior work is that there are inconsistencies in how emotions are categorised; whether or not items are considered uniquely human changes between studies. For example, ‘enjoyment’ is considered uniquely human whereas ‘joy’ is not (Leyens et al., 2001; Paladino et al., 2002). ‘Happiness’ is sometimes considered uniquely human (Paladino et al., 2002) but sometimes not (Cortes et al., 2005). These problems may arise because previous studies have not rigorously pretested the emotion terms they used as stimuli as well as from translational discrepancies across studies conducted in different languages. In order to formally determine which emotions (In English) tend to be considered uniquely human and which tend to be considered shared with other animals, we conducted a pretest in which we asked participants to rate fifty-four common emotion terms on three scales: human uniqueness, valence of experience and sociality.

3.1. Pretest Methods

3.1.1. Participants

Sixty participants completed the ratings (22 female, 37 male, 1 ‘other’), aged between 18 and 54 (Mean age=26.8, *SD*=7.98). All participants were fluent in English. Eight people failed one or more attention checks and their data was excluded and replaced.

3.1.2. Scales

We chose fifty-four common emotion terms and asked participants to rate them on Humanness (the extent to which it is believed each emotion is experienced by humans compared to other species), Valence (the extent to which it is believed each emotion is positive or negative to experience) and Sociality (the extent to which it is

241 believed each emotion is prosocial or antisocial), using three separate sliding
242 scales. The full list of emotion terms, along with additional information about scale
243 presentation, is in supplementary information.

244 **3.1.3. Procedure**

245 Participants were informed that the study would examine the ways in which people
246 understand emotional terms and that they would be asked to rate emotion words on
247 the three separate scales. Once informed consent was obtained, brief demographic
248 and screening questions were asked. Then, participants were taken through the
249 three question blocks. Participants were debriefed and redirected back to Prolific for
250 payment. The session took approximately twelve minutes.

251 **3.2. Pretest results and discussion**

252 Our pretest confirmed that emotions differ both in valence and sociality. We present
253 the mean ratings for each emotion on Humanness, Valence and Sociality in
254 supplementary information, Table S1. The 'basic' (or 'primary') emotions such as
255 fear, sadness, happiness and surprise featured among the emotions most thought
256 of as shared with other species (Ekman, 1992). In line with infrahumanization
257 theory, we largely replicated prior work from Demoulin et al. (2004), who also
258 reported emotions such as nostalgia and optimism to be most uniquely human, and
259 emotions such as fear and surprise to be least uniquely human. Importantly,
260 however, none of the terms commonly included as negative secondary emotions in
261 previous research, such as guilt, remorse, resignation and melancholy (Leyens et
262 al., 2001; Paladino et al., 2002) were perceived to be antisocial, even though they
263 were considered negative to experience.

Overall, there was a general positive association between mean scores across participants for valence and sociality. This suggests that, across a broad range of emotion terms, emotions that make us feel positive are also viewed as prosocial and emotions that make us feel negative may be viewed as more antisocial. However, for the specific negative emotions commonly included in prior infrahumanization work, participants' scores on valence and sociality scales were not strongly (if at all) associated. For example, correlations were $r(58) = .131$, $p = .318$ for regret, $r(58) = .187$, $p = .153$ for melancholy, $r(58) = .262$, $p = .035$ for disillusion, and $r(58) = .060$, $p = .651$ for remorse. This shows that the kinds of negative emotions that infrahumanization researchers have included in previous research (e.g., regret, melancholy, disillusion, remorse - Banton et al., 2020; Leyens et al., 2001; Paladino et al., 2002; Vaes et al., 2003) may be negative to experience but are not antisocial and so are not best placed to determine whether infrahumanization is separate from intergroup preference. This provides strong grounds for re-examining the nature of intergroup bias in emotion attribution.

4. Study 1: Replicating previous research

In our first six experiments we sought to replicate previous research. Participants rated how strongly they believed ingroup and outgroup members to experience sixteen emotions. Four emotions were unique to humans and positive (nostalgia, optimism, humility, hope), four were unique to humans and negative, (disillusion, regret, melancholy, remorse) four were shared with other animals and positive, (happiness, tenderness, surprise, love) and four were shared with other animals and negative (fear, loneliness, sadness, nervousness).

4.1. Study 1 Methods

4.1.1. Participants

A power analysis using MorePower 6.0.4 found a minimum N of 126 to be necessary to detect interactions with a medium effect size (partial eta squared .06) with an alpha of .05 and power of .8. 130 different participants completed the ratings in each experiment. Participants were only eligible if they were 18 or over, fluent in English and had not taken part in any of the other experiments reported. Data collection for each experiment took place completely separately. We excluded and replaced any participants that failed one or more of the attention checks.

In **Experiment 1a** (Muslim outgroup), participants could only take part if they identified as Christian. Seven people failed one or more attention check. Of the final sample, 95 participants were female and 35 were male, aged from 18 to 68 (Mean age = 35.3, *SD* = 14.07).

In **Experiment 1b** (criminal outgroup), participants could only take part if they had not served previous prison sentences. Five people failed one or more attention check. Of the final sample, 58 participants were female, 70 were male, 1 was nonbinary and 1 indicated 'prefer not to say'. Ages ranged from 18 to 59 (Mean age = 27.1, *SD* = 7.94).

In **Experiment 1c** (child molester outgroup), four people failed one or more attention check. Of the final sample, 84 participants were female, 46 were male, and ages ranged from 18 to 57 (Mean age = 28.5, *SD* = 10.66).

In **Experiment 1d** (anti-vaxxer outgroup), participants could only take part if they were pro vaccination. Four people failed one or more attention check and three additional people were excluded because they indicated they were anti vaccination.

Of the final sample, 53 participants were female, 75 were male and 2 were non-binary/agender, aged from 18 to 60 (Mean age = 27.0, *SD* = 8.92).

In **Experiment 1e** (non-social distancer outgroup), participants could only take part if they were living in the UK and reported that they were following current social distancing regulations. Three people were excluded because they indicated that they were not adhering to social distancing regulations and their data was excluded and replaced. Of the final sample, 95 participants were female and 35 were male, aged from 18 to 68 (Mean age = 35.3, *SD* = 14.07).

In **Experiment 1f** (climate change deniers as outgroup), participants could only take part if they believed in climate change. Five people failed one or more attention check. Of the final sample, 52 participants were female, 78 were male and ages ranged from 18 to 62 (Mean age = 26.6, *SD* = 9.22).

4.1.2. Stimuli development

We chose emotions from our pretest data (supplementary information, Table S1) that best fit the four emotion categories of interest: unique to humans and positive, unique to humans and negative, shared with other animals and positive and shared with other animals and negative. Table 1 shows the list of emotion words included in the final stimulus sets. From the most and least uniquely human terms, we chose four rated as highly positive and four rated as highly negative. In developing the items for our emotion categories, we ensured that humanness ratings were closely matched between the positive and negative conditions for each level of humanness so that we could accurately separate valence effects from ones of humanness. See supplementary information for further details on stimuli development.

Table 1. Emotion terms included for each condition in all experiments

	Study 1		Study 2	
	Positive	Negative	Prosocial	Antisocial
Unique to humans	Nostalgia	Disillusion	Nostalgia	Arrogance
	Optimism	Regret	Optimism	Schadenfreude
	Humility	Melancholy	Humility	Contempt
	Hope	Remorse	Hope	Scorn
Shared with other animals	Happiness	Fear	Happiness	Hostility
	Tenderness	Loneliness	Tenderness	Irritation
	Surprise	Sadness	Love	Anger
	Love	Nervousness	Friendliness	Disgust

4.1.3. Scales

Participants indicated on unmarked sliders how strongly they thought the ingroup and outgroup in each experiment experienced the sixteen emotions from *Not at all* (0) to *Very strongly* (100), with the midpoint *Somewhat* (50). For example, in Experiment 1b, the outgroup block began 'In the following questions, please consider the group: **Individuals with criminal convictions'**. Then, participants would respond to each item, such as 'How strongly do you think a typical criminal feels **nostalgia**'. Ingroup and outgroup items were presented in two separate blocks shown on sequential screens, the order of which was counterbalanced across participants. The sixteen emotion items within each block were randomised and one attention check per block was also included approximately halfway through, such as 'Please indicate *Somewhat'.*

Participants also completed the blatant dehumanization scale (Kteily et al., 2015) (Figure 1) and a simple preference measure for both groups. In the blatant dehumanization scale, participants saw the 'ascent of man' image and were asked to indicate on an unmarked slider how evolved they considered the average

member of each group to be, with 0 corresponding to the very bottom and 100 to the most human at the very top. In the attitude scale, participants were asked to indicate how they felt about each group using an unmarked sliding scale from *Extremely Negative* (0) to *Extremely Positive* (100). For all scales, half of the participants responded to ingroup items first and half to outgroup items first.

We included the group preference and blatant dehumanization measures to check that our chosen groups were the kinds that we should expect to see infrahumanized should the process occur. Prior work shows that infrahumanization measures correlate positively with blatant dehumanization scores (Kteily et al., 2015). Thus, though they are not claimed to measure the same construct, they have been shown to reliably co-occur. We included the attitude measure as confirmation that the outgroups were social categories that participants 'would not like or want to belong to' (Leyens, 2009), also increasing chances of detecting infrahumanization if it occurs.

4.1.4. Procedure

Participants were informed that the study was designed to help us understand the ways in which people ascribe emotions to different groups of individuals and stated the particular groups of interest for each experiment. Participants were instructed that they would be asked to rate sixteen emotion words on two scales, one for each social category, and then complete two scales asking about attitudes to each group. Once informed consent was obtained, brief demographic and screening (if relevant) questions were asked. Then, participants were taken through the two experimental 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 sessions took under ten minutes.

4.1.5. Design and data analysis

In line with our pre-registered analysis plan, we conducted 2 (Group: ingroup/outgroup) x 2 (Valence: positive/negative) x 2 (Humanness: unique to humans / shared with other animals) within subjects ANOVAs to test for infrahumanization in intergroup emotion attributions. Scores for each emotion category were obtained by calculating the mean of the four emotion terms within the category for each participant. For example, a participant's score for uniquely human positive emotion ascriptions towards the ingroup would be the mean of their ratings on Nostalgia, Optimism, Humility and Hope within the ingroup block. More detail about the design is available in supplementary information.

In this design, infrahumanization would be observed in an interaction whereby uniquely human emotions are more strongly ascribed to the ingroup, independent of valence (Leyens et al., 2000). This should not be the case for emotions shared with other animals, for which previous work found the reverse or no difference (Leyens et al., 2001). For example, in original experiments demonstrating infrahumanization, Leyens and colleagues (2001) showed that more positive and negative uniquely human emotions were attributed to the ingroup than the outgroup and that this was not qualified by an interaction with valence (see Leyens et al., 2001, p. 402, Fig. 2).

Though previous studies do not find interactions between intergroup emotion attributions and valence, there tend to be main effects of valence such that ascriptions of positive emotions are generally higher than negative ones. These

results are not central to the predictions of infrahumanization theory nor the social preference account.

In following up significant interactions, we report only comparisons between ingroup and outgroup ratings for each condition, in line with testing the main hypotheses. We measured differences in ratings for ingroup and outgroup on the attitude and 'blatant dehumanization' scales using paired-samples t-tests. All tests were two-sided and met the assumptions necessary for our statistical approaches.

4.2. Study 1 Results

4.2.1. Blatant dehumanization and attitude scores

In every experiment, the outgroup was rated as significantly less human than the ingroup on the blatant dehumanization scale (all *p*s<.001). Additionally, participants reported feeling significantly more negative towards the outgroup than the ingroup on the attitude scale (all *p*s<.001). Figure 1 shows the points at which outgroups and ingroups were marked on the blatant dehumanization scale. The extent to which outgroups were 'blatantly dehumanized' varied greatly across our intergroup contexts. The average point at which Muslims and 'criminals' were marked fell between the most 'evolved' looking human silhouette and the more caveman-like silhouette next to it on the scale. 'Child molesters' and 'non-social distancers' were rated much further down on the ascent scale, nearer to the midway point between the ape-like and modern human-like depictions. Figure S1 (supplementary information) shows mean results for each ingroup and outgroup on the dehumanization and attitude measures.

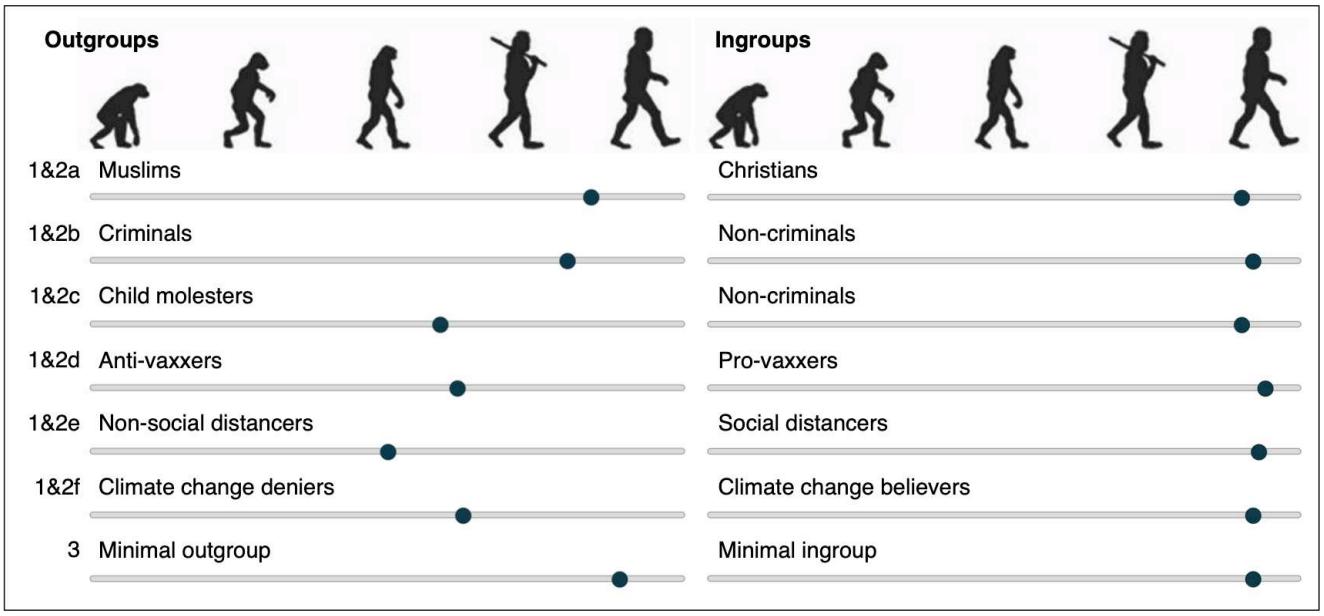

Figure 1. The average points at which outgroups and ingroups were marked on the blatant dehumanization scale across Studies 1, 2 and 3. All outgroups were significantly dehumanized relative to the corresponding ingroup (all *ps*<.001 in Studies 1 &2, $p = .002$ in Study 3).

4.2.2. Intergroup emotion ascriptions

4.2.2.1. Experiment 1a

For ratings towards Muslims (outgroup) and Christians (ingroup), there were significant main effects of group, $F(1, 129) = 32.54$, $p < .001$, $\eta_p^2 = .201$, and of valence, $F(1, 129) = 101.80$, $p < .001$, $\eta_p^2 = .441$, with ratings higher overall for ingroup than outgroup and for positive than negative emotions. There was no significant main effect of humanness, $F(1, 129) = .76$, $p = .384$, $\eta_p^2 = .006$.

There were significant interactions between group and humanness, $F(1, 129) = 11.89$, $p = .001$, $\eta_p^2 = .084$, group and valence, $F(1, 129) = 17.16$, $p < .001$, $\eta_p^2 = .117$, and humanness and valence, $F(1, 129) = 11.77$, $p = .001$, $\eta_p^2 = .084$. Pairwise comparisons showed that overall, ratings were higher for ingroup than outgroup

both for uniquely human emotions (p<.001) and for emotions shared with other animals (p = .001). Ratings were overall higher for ingroup than outgroup for positive emotions (p<.001), but there were no differences between groups for negative emotions (p = .463).

All effects were qualified in a significant three-way interaction, $F(1, 129)$ = 11.37, p =.001, η_p^2= .081. Planned comparisons showed that ratings were higher for ingroup than outgroup for positive uniquely human terms (p < .001), negative uniquely human terms (p =.006), and positive terms shared with other animals (p <.001). However, there was no difference between ingroup and outgroup for negative terms shared with other animals (p = .104).

4.2.2.2. Experiment 1b

For ratings towards Criminals (outgroup) and 'individuals with no criminal history' (ingroup), there were significant main effects of group, $F(1, 129) = 31.84$, p <.001, η_p^2= .198, valence, $F(1, 129) = 4.64$, p =.033, η_p^2= .035, and humanness, $F(1, 129)$ = 35.86, p <.001, η_p^2= .218. Ratings were higher overall for ingroup than outgroup, for negative than positive emotions, and for emotions shared with other animals than for uniquely human emotions.

There were significant interactions between group and humanness, $F(1, 129)$ = 7.62, p =.007, η_p^2= .056, group and valence, $F(1, 129) = 167.70$, p <.001, η_p^2= .565, and humanness and valence, $F(1, 129) = 4.16$, p =.043, η_p^2= .031. Pairwise comparisons showed that overall, ratings were higher for ingroup than outgroup both for uniquely human emotions and for emotions shared with other animals (ps < .001). Ratings were overall higher for ingroup than outgroup for positive emotions (p<.001), and higher for outgroup than ingroup for negative emotions (p = .006).

All effects were qualified in a significant three-way interaction, $F(1, 129) = 4.56$, $p = .035$, $\eta_p^2 = .034$. Planned comparisons showed that ratings were higher for ingroup than outgroup for positive terms, both uniquely human and shared with other animals ($ps < .001$), but higher for outgroup than ingroup on negative terms, both uniquely human ($p = .007$) and shared with other animals ($p = .020$).

4.2.2.3. Experiment 1c

For ratings towards 'child molesters' (outgroup) and 'individuals with no criminal history' (ingroup), there were significant main effects of group, $F(1, 129) = 154.31$, $p < .001$, $\eta_p^2 = .545$, and of humanness, $F(1, 129) = 83.51$, $p < .001$, $\eta_p^2 = .393$, but not of valence, $F(1, 129) = 2.97$, $p = .087$, $\eta_p^2 = .023$. Ratings were higher overall for ingroup than outgroup and for emotions shared with other animals than for uniquely human emotions.

There were significant interactions between group and humanness, $F(1, 129) = 17.70$, $p < .001$, $\eta_p^2 = .121$, group and valence, $F(1, 129) = 119.32$, $p < .001$, $\eta_p^2 = .481$, and humanness and valence, $F(1, 129) = 28.98$, $p < .001$, $\eta_p^2 = .189$. Pairwise comparisons showed that overall, ratings were higher for ingroup than outgroup both for uniquely human emotions and for emotions shared with other animals ($ps < .001$). Ratings were also overall higher for ingroup than outgroup for positive emotions ($p < .001$), and for negative emotions ($p = .001$).

All effects were qualified in a significant three-way interaction, $F(1, 129) = 14.10$, $p < .001$, $\eta_p^2 = .099$. Planned comparisons showed that ratings were higher for ingroup than outgroup for positive uniquely human terms ($p < .001$), negative uniquely human terms ($p < .001$), and positive terms shared with other animals (p

<.001). However, there was no difference between ingroup and outgroup for negative terms shared with other animals (p = .287).

4.2.2.4. Experiment 1d

For ratings towards 'anti-vaxxers' (outgroup) and 'pro-vaxxers' (ingroup), there were significant main effects of humanness, $F(1, 129) = 40.42$, $p < .001$, $\eta_p^2 = .239$, and of valence, $F(1, 129) = 69.59$, $p < .001$, $\eta_p^2 = .350$, but not of group, $F(1, 129) = 1.02$, $p = .315$, $\eta_p^2 = .008$. Ratings were higher overall emotions shared with other animals than for uniquely human emotions and for positive than negative emotions.

There were significant interactions between Group and Valence, $F(1, 129) = 88.99$, $p < .001$, $\eta_p^2 = .408$, Group and Humanness, $F(1, 129) = 11.49$, $p = .001$, $\eta_p^2 = .082$, and Valence and Humanness, $F(1, 129) = 8.41$, $p = .004$, $\eta_p^2 = .061$. Pairwise comparisons showed that overall, ratings were higher for ingroup than outgroup for uniquely human emotions (p = .017) but not for emotions shared with other animals ($p < .358$). Ratings were overall higher for ingroup than outgroup for positive emotions (p<.001) but higher for outgroup than ingroup for negative emotions (p = .001). The three-way interaction was not significant, $F(1, 129) = .31$, $p = .580$, $\eta_p^2 = .002$.

4.2.2.5. Experiment 1e

For ratings towards 'non-social distancers' (outgroup) and 'social distancers' (ingroup), there were significant main effects of group, $F(1, 129) = 239.50$, $p < .001$, $\eta_p^2 = .650$, and of humanness, $F(1, 129) = 60.13$, $p < .001$, $\eta_p^2 = .318$, but not of valence, $F(1, 129) = 1.75$, $p = .188$, $\eta_p^2 = .013$. Ratings were higher overall for ingroup than outgroup and for emotions shared with other animals than for uniquely human emotions.

There were significant interactions between group and humanness, $F(1, 129) = 38.46$, $p < .001$, $\eta_p^2 = .230$, group and valence, $F(1, 129) = 16.08$, $p < .001$, $\eta_p^2 = .111$, and humanness and valence, $F(1, 129) = 113.12$, $p < .001$, $\eta_p^2 = .467$. Pairwise comparisons showed that overall, ratings were higher for ingroup than outgroup both for uniquely human emotions and for emotions shared with other animals (*ps* < .001). Ratings were also overall higher for ingroup than outgroup for positive emotions ($p < .001$), and for negative emotions ($p = .001$).

All effects were qualified in a significant three-way interaction, $F(1, 129) = 149.18$, $p < .001$, $\eta_p^2 = .536$. Planned comparisons showed that ratings were higher for ingroup than outgroup for all four emotion conditions - positive uniquely human terms, negative uniquely human terms, positive terms shared with other animals and negative terms shared with other animals (all *ps* <.001).

4.2.2.6. Experiment 1f

For 'climate change deniers' (outgroup) and 'climate change believers' (ingroup), there were significant main effects of group, $F(1, 129) = 171.51$, $p < .001$, $\eta_p^2 = .571$, and of humanness, $F(1, 129) = 27.79$, $p < .001$, $\eta_p^2 = .177$, but not of valence, $F(1, 129) = .85$, $p = .359$, $\eta_p^2 = .007$. Ratings were higher overall for ingroup than outgroup and for emotions shared with other animals than for uniquely human emotions.

There were significant interactions between group and humanness, $F(1, 129) = 3.92$, $p = .05$, $\eta_p^2 = .029$, group and valence, $F(1, 129) = 38.99$, $p < .001$, $\eta_p^2 = .232$, and humanness and valence, $F(1, 129) = 17.02$, $p < .001$, $\eta_p^2 = .117$. Pairwise comparisons showed that overall, ratings were higher for ingroup than outgroup both for uniquely human emotions and for emotions shared with other animals (*ps* < .001). Ratings were also overall higher for ingroup than outgroup for positive

emotions and for negative emotions (*ps* = .001). The three-way interaction was not significant, $F(1, 129) = .13$, $p = .724$, $\eta_p^2 = .001$.

Mean scores (M) and standard errors of the mean (SE) for each of the conditions across all experiments in Study 1 are shown in supplementary information, Table S2. Figure 2 shows results from Study 1.

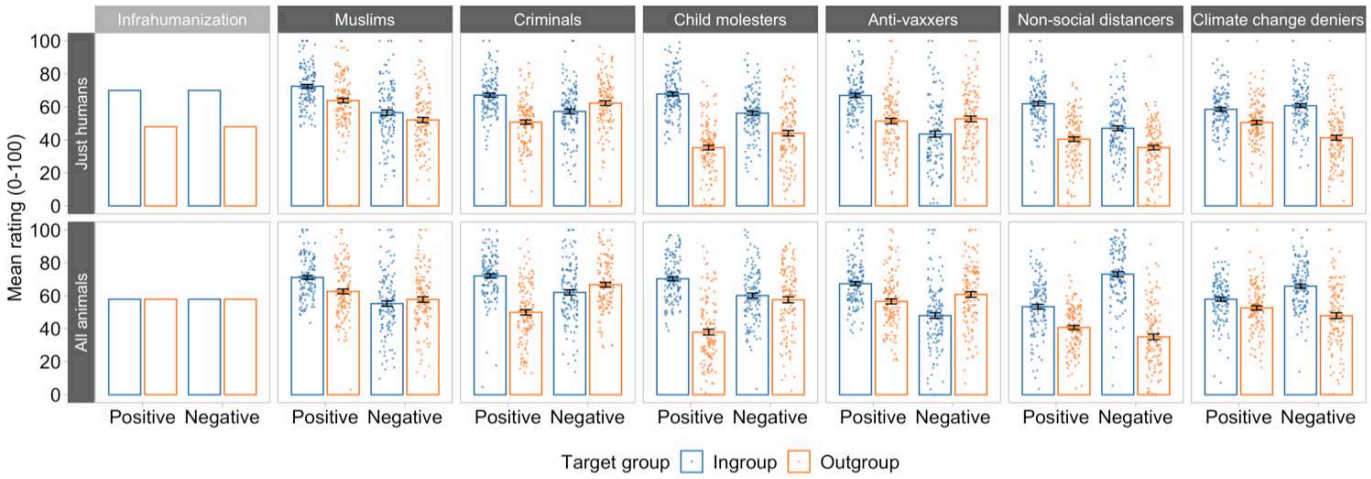

Figure 2. Partial replications of infrahumanization theory. Some outgroups (Muslims, child molesters, non-social distancers and climate change deniers) were rated overall as experiencing both positive and negative uniquely human emotions to a lesser extent than the ingroup (top panels). These outgroups were also rated as experiencing some emotions shared with other animals to a lesser extent than the ingroup (bottom panels). Note that while main effects of valence have been reported in some prior work (see Leyens et al., 2001, p. 402, Fig. 2), we do not plot these the prediction, as they are not relevant for distinguishing between the theories. Error bars represent standard errors.

4.3. Study 1 Discussion

Results partially replicated the predictions of infrahumanization theory - some outgroups were rated overall as experiencing both positive and negative uniquely

553 human emotions to a lesser extent than the ingroup. However, these outgroups
554 were also rated as experiencing emotions shared with other animals to a lesser
555 extent than the ingroup. In each experiment, the outgroup was rated as 'less human'
556 on the blatant dehumanization scale, confirming that these were the types of
557 intergroup contexts in which we ought to see infrahumanization effects should the
558 process occur. We next examined whether controlling for the sociality of emotional
559 terms explained apparent evidence for infrahumanization.
560
561 **5. Study 2: Testing the social preference account**
562 In our next six experiments, we test whether what appears to be infrahumanization
563 may be explained by ingroup preference and stereotyping. Rather than comparing
564 intergroup ascriptions of emotions that varied by how positive or negative they are to
565 experience, we compared ascriptions of emotions than varied by how prosocial or
566 antisocial they are in character. Participants rated the same six groups on four types
567 of emotional experience: unique to humans and prosocial (nostalgia, optimism,
568 humility, hope), unique to humans and antisocial (arrogance, schadenfreude,
569 contempt, scorn), shared with other animals and prosocial (happiness, tenderness,
570 love, friendliness) and shared with other animals and antisocial (hostility, irritation,
571 anger, disgust).
572 This design pits the predictions of infrahumanization against a social
573 preference account. Infrahumanization would be observed in an interaction between
574 Group and Humanness such that uniquely human emotions will be more strongly
575 ascribed to ingroup than outgroup, both for prosocial and antisocial emotions (i.e.,
576 Leyens et al., 2001, p. 402, Fig. 2). In contrast, we hypothesised that prosocial

emotions will typically be attributed more strongly to the ingroup and antisocial ones to the outgroup, regardless of humanness (Figure 3 shows both predictions).

5.1. Study 2 Methods

5.1.1. Participants

Based on the same power analysis as reported for Study 1, 130 different participants completed the ratings in each experiment. The same eligibility criteria were applied as for Study 1.

For **Experiment 2a** (Muslim outgroup), participants could only take part if they identified as Christian. Six people failed one or more attention check. Of the final sample, 86 participants were female, 42 male and 2 were non-binary/agender, aged from 18 to 71 (Mean age = 33.6, *SD* = 11.79).

For Experiment 2b (criminal outgroup), participants could only take part if they had not served previous prison sentences. One person failed one or more attention check. Of the final sample, 62 participants were female, 68 were male and ages ranged from 18 to 65 (Mean age = 26.9, *SD* = 8.65).

For **Experiment 2c** (child molester outgroup), three people failed one or more attention check. Of the final sample, 87 were female, 42 male and 1 non-binary, with an age range of 18 to 61 (Mean age = 31.6, *SD* = 10.14).

In **Experiment 2d** (anti-vaxxer outgroup), participants could only take part if they were pro-vaccination. Seven people failed one or more attention check and four additional people were excluded because they indicated that they were anti vaccination. Of the final sample, 50 were female and 80 were male, aged from 18 to 51 (Mean age = 25.9, *SD* = 8.01).

In **Experiment 2e** (non-social distancers outgroup), participants could only take part if they were living in the UK and following social distancing regulations. Data was excluded and replaced for three participants who failed one or more attention check and one additional participant who indicated that they were not adhering to social distancing regulations. Of the final sample, 86 were female, 42 were male and 2 were non-binary/agender, aged from 18 to 71 (Mean age = 33.6, *SD* = 11.79).

For **Experiment 2f** (climate change deniers outgroup), participants could only take part if they believed in climate change. Two people failed one or more attention check. Of the final sample, 53 were female, 76 male and 1 non-binary, with an age range of 18 to 60 (Mean age = 27.3, *SD* = 8.36).

5.1.2. Stimuli development

We chose emotions from our pretest data (supplementary information, Table S1) that best fit the four emotion categories of interest: unique to humans and prosocial, unique to humans and antisocial, shared with other animals and prosocial and shared with other animals and antisocial. From the most and least uniquely human terms, we chose four rated as highly prosocial and four rated as highly antisocial, this time ignoring valence ratings. Table 1 shows the full list of emotion words. We chose the emotions such that humanness ratings were closely matched between the prosocial and antisocial conditions at each level of humanness. This was so that dimensions of Sociality and Humanness were orthogonal, allowing us to accurately separate effects of each. See supplementary information for further details on stimuli development.

5.1.3. Scales

We employed the same six intergroup contexts as for Study 1. Apart from including different emotion items, the emotion attribution scales were identical as to those described for Study 1. Participants again completed the blatant dehumanization scale (Kteily et al., 2015) (Figure 1) and the group preference scale for the ingroup and outgroup in each experiment.

5.1.4. Procedure, design and data analysis

The procedure was the same as outlined for Study 1. The design and data analysis were almost identical as described for Study 1 though with the Sociality (prosocial/antisocial) variable instead of the Valence (positive/negative) variable.

5.2. Study 2 Results

5.2.1. Blatant dehumanization and attitude scores

The outgroup was always rated as significantly less human than the ingroup on the blatant dehumanization scale (all *p*s<.001). Additionally, participants reported feeling significantly more negative towards the outgroup than the ingroup on the attitude scale (all *p*s<.001) (Figure 1). Figure S1 (supplementary information) shows mean results for each ingroup and outgroup on the dehumanization and attitude measures.

5.2.2. Intergroup emotion ascription ratings

5.2.2.1. Experiment 2a

For ratings towards Muslims (outgroup) and Christians (ingroup), there were main effects of humanness, $F(1, 129) = 28.75$, $p < .001$, $\eta_p^2 = .182$, and of sociality, $F(1, 129) = 147.39$, $p < .001$, $\eta_p^2 = .533$, but not of group, $F(1, 129) = .42$, $p = .517$, $\eta_p^2 = .003$. Ratings were higher overall for emotions shared with other animals than for uniquely human emotions, and for prosocial than antisocial emotions.

There was a significant interaction between group and sociality, $F(1, 129) = 39.45$, $p < .001$, $\eta_p^2 = .234$, but not between group and humanness, $F(1, 129) = .75$, $p = .389$, $\eta_p^2 = .006$, nor between humanness and sociality, $F(1, 129) = 1.74$, $p = .190$, $\eta_p^2 = .013$. Pairwise comparisons showed that overall, ratings were higher for ingroup than outgroup for prosocial emotions, but higher for outgroup than ingroup for antisocial emotions (*ps* < .001).

All effects were qualified in a significant three-way interaction $F(1, 129) = 3.97$, $p = .048$, $\eta_p^2 = .030$. Planned comparisons showed that ratings were higher for ingroup than outgroup on prosocial terms, both uniquely human and shared with other animals (*ps* < .001), and higher for outgroup than ingroup on antisocial terms (*ps* < .001), both uniquely human and shared with other animals.

5.2.2.2. Experiment 2b

For ratings towards 'convicted criminals' (outgroup) and 'individuals with no criminal history' (ingroup), there were main effects of humanness, $F(1, 129) = 40.04$, $p < .001$, $\eta_p^2 = .237$, and of group, $F(1, 129) = 36.63$, $p < .001$, $\eta_p^2 = .221$, but not of sociality, $F(1, 129) < .001$, $p = .996$, $\eta_p^2 < .001$. Ratings were higher overall for emotions shared with other animals than for uniquely human emotions, and for ingroup than outgroup.

There was a significant interaction between group and sociality, $F(1, 129) = 201.29$, $p < .001$, $\eta_p^2 = .609$, humanness and sociality, $F(1, 129) = 33.63$, $p < .001$, $\eta_p^2 = .207$, but not between group and humanness, $F(1, 129) = .98$, $p = .323$, $\eta_p^2 = .008$. Pairwise comparisons showed that overall, ratings were higher for ingroup than outgroup for prosocial emotions, but higher for outgroup than ingroup for antisocial emotions (*ps* < .001).

All effects were qualified in a significant three-way interaction, $F(1, 129) = 24.72$, $p < .001$, $\eta_p^2 = .161$. Planned analyses of simple effects following the three-way interaction showed that ratings were higher for ingroup than outgroup on prosocial terms, both uniquely human and shared with other animals ($ps < .001$), and higher for outgroup than ingroup on antisocial terms ($ps < .001$), both uniquely human and shared with other animals.

5.2.2.3. Experiment 2c

For ratings towards 'child molesters' (outgroup) and 'individuals with no criminal history' (ingroup) on emotion experiences, there were main effects of humanness, $F(1, 129) = 6.81$, $p = .010$, $\eta_p^2 = .050$, and of group, $F(1, 129) = 122.42$, $p < .001$, $\eta_p^2 = .487$, and of sociality, $F(1, 129) = 25.01$ $p < .001$, $\eta_p^2 = .162$. Ratings were higher overall for emotions shared with other animals than for uniquely human emotions, for ingroup than outgroup, and for antisocial than prosocial emotions.

There was a significant interaction between group and sociality, $F(1, 129) = 201.29$, $p < .001$, $\eta_p^2 = .609$, humanness and sociality, $F(1, 129) = 234.42$, $p < .001$, $\eta_p^2 = .645$, and between group and humanness, $F(1, 129) = 21.82$, $p < .001$, $\eta_p^2 = .145$. Pairwise comparisons showed that overall, ratings were higher for ingroup than outgroup for prosocial emotions, but higher for outgroup than ingroup for antisocial emotions ($ps < .001$). Though the interaction between group and humanness was significant, this did not reflect the infrahumanization prediction. Ratings were overall higher for ingroup than outgroup both for uniquely human emotions and for emotions shared with other animals. Importantly, ratings were higher for ingroup than outgroup on prosocial terms, both uniquely human and shared with other animals ($ps. < .001$), and higher for outgroup than ingroup on

antisocial terms, both uniquely human (*p.* <.001) and shared with other animals (*p* = .004). The three-way interaction was not significant, $F(1, 129) = .28$, $p = .600$, $\eta_p^2 = .002$.

5.2.2.4. Experiment 2d

For ratings towards 'anti-vaxxers' (outgroup) and 'pro-vaxxers' (ingroup), there was a significant main effect of humanness, $F(1, 129) = 69.28$, $p < .001$, $\eta_p^2 = .349$, but not of group, $F(1, 129) = 2.79$, $p = .097$, $\eta_p^2 = .021$, nor of sociality, $F(1, 129) = 1.33$, $p = .251$, $\eta_p^2 = .010$. Ratings were higher overall for emotions shared with other animals than for uniquely human emotions.

There was a significant interaction between group and sociality, $F(1, 129) = 216.29$, $p < .001$, $\eta_p^2 = .626$, humanness and sociality, $F(1, 129) = 4.55$, $p = .035$, $\eta_p^2 = .034$, but not between group and humanness, $F(1, 129) = .08$, $p = .782$, $\eta_p^2 = .001$. Pairwise comparisons showed that overall, ratings were higher for ingroup than outgroup for prosocial emotions, but higher for outgroup than ingroup for antisocial emotions (*ps* < .001).

All effects were qualified in a significant three-way interaction $F(1, 129) = 14.83$, $p < .001$, $\eta_p^2 = .103$. Planned comparisons showed that ratings were higher for ingroup than outgroup on prosocial terms, both uniquely human and shared with other animals and higher for outgroup than ingroup on antisocial terms both uniquely human and shared with other animals (all *ps* <.001).

5.2.2.5. Experiment 2e

For ratings towards 'non-social distancers' (outgroup) and 'social distancers' (ingroup), there was a significant main effect of humanness, $F(1, 129) = 10.32$, $p = .002$, $\eta_p^2 = .074$, but not of group, $F(1, 129) = .30$, $p = .584$, $\eta_p^2 = .002$, nor of

720 sociality, $F(1, 129) = 1.80$, $p = .183$, $\eta_p^2 = .014$. Ratings were higher overall for
721 emotions shared with other animals than for uniquely human emotions.
722 There was a significant interaction between group and sociality, $F(1, 129) =$
723 213.36, $p < .001$, $\eta_p^2 = .623$, group and humanness, $F(1, 129) = 59.99$, $p < .001$, $\eta_p^2 =$
724 .306, and sociality and humanness, $F(1, 129) = 56.59$, $p < .001$, $\eta_p^2 = .305$. Pairwise
725 comparisons showed that overall, ratings were higher for ingroup than outgroup for
726 prosocial emotions, but higher for outgroup than ingroup for antisocial emotions (*ps*
727 < .001). Ratings were also higher for outgroup than ingroup for uniquely human
728 emotions, but higher for ingroup than outgroup on emotions shared with other
729 animals (*ps* < .001).
730 All effects were qualified in a significant three-way interaction $F(1, 129) =$
731 37.90, $p < .001$, $\eta_p^2 = .227$. Planned comparisons showed that ratings were higher for
732 ingroup than outgroup on prosocial terms, both uniquely human and shared with
733 other animals and higher for outgroup than ingroup on antisocial terms both uniquely
734 human and shared with other animals (all *ps* <.001).
735 **5.2.2.6.Experiment 2f**
736 For 'climate change deniers' (outgroup) and 'climate change believers' (ingroup),
737 there was a significant main effect of humanness, $F(1, 129) = 102.37$, $p < .001$, $\eta_p^2 =$
738 .442, of group, $F(1, 129) = 6.32$, $p = .013$, $\eta_p^2 = .047$, and a marginal effect of
739 sociality, $F(1, 129) = 3.85$, $p = .052$, $\eta_p^2 = .029$. Ratings were higher overall for
740 emotions shared with other animals than for uniquely human emotions, for ingroup
741 than outgroup, and for antisocial than prosocial emotions.
742 There was a significant interaction between group and sociality, $F(1, 129) =$
743 47.03, $p < .001$, $\eta_p^2 = .267$, group and humanness, $F(1, 129) = 84.72$, $p < .001$, $\eta_p^2 =$

744 .396, and sociality and humanness, $F(1, 129) = 9.37$, $p = .003$, $\eta_p^2 = .068$. Pairwise
745 comparisons showed that overall, ratings were higher for ingroup than outgroup for
746 prosocial emotions, but higher for outgroup than ingroup for antisocial emotions (*ps*
747 < .001). Ratings were also higher for outgroup than ingroup for uniquely human
748 emotions, but higher for ingroup than outgroup on emotions shared with other
749 animals (*ps* < .001).
750 All effects were qualified in a significant three-way interaction $F(1, 129) =$
751 14.04, $p < .001$, $\eta_p^2 = .098$. Planned comparisons showed that ratings were higher for
752 ingroup than outgroup on prosocial terms, both uniquely human ($p = .001$) and
753 shared with other animals ($p < .001$), and higher for outgroup than ingroup on
754 uniquely human antisocial terms ($p < .001$). However, there was no difference
755 between ingroup and outgroup on antisocial terms shared with other animals ($p =$
756 .200).
757 Mean scores (M) and standard errors of the mean (SE) for each of the
758 conditions across all experiments in Study 2 are shown in supplementary
759 information, Table S3. Figure 3 shows results for Study 2.

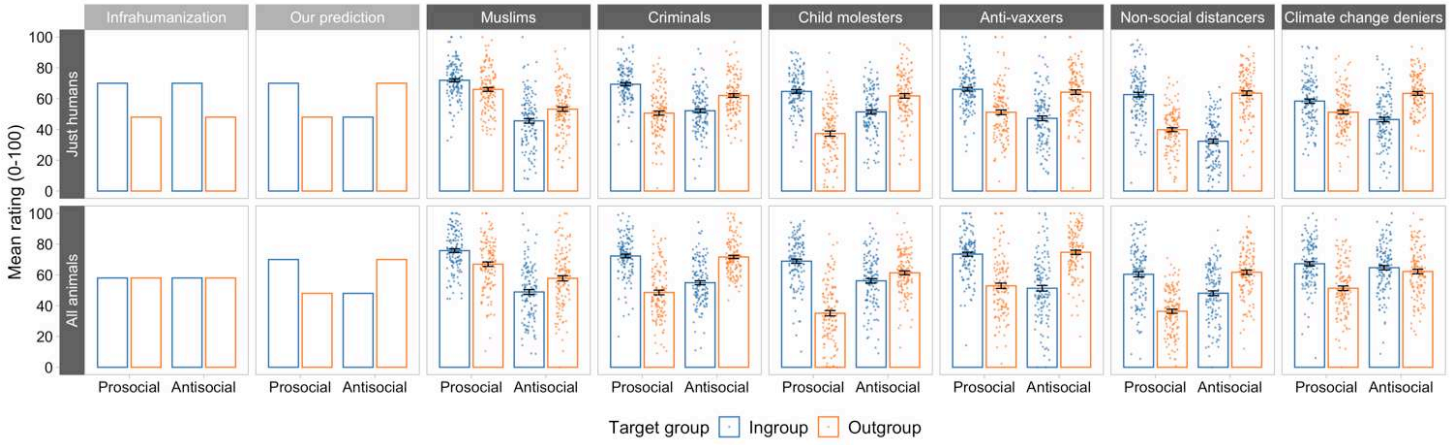

Figure 3. Evidence for social preferences but not infrahumanization. Contrary to infrahumanization theory, ratings were higher for the ingroup than the outgroup for

prosocial emotions but higher for the outgroup than the ingroup for antisocial emotions across all group contexts. Note that while main effects of valence have been reported in some prior work (see Leyens et al., 2001, p. 402, Fig. 2), we do not plot these in our predictions, as they are not relevant for distinguishing between the theories. Error bars represent standard errors.

5.3. Study 2 Discussion

Contrary to infrahumanization theory, outgroups were not denied uniquely human emotions relative to ingroups. All outgroups were thought to experience prosocial emotions to a lesser extent than ingroups, both for uniquely human emotions and for those shared with other animals. However, all outgroups were also thought to experience uniquely human antisocial emotions to a greater extent than ingroup members. Muslims, criminals, child molesters, anti-vaxxers and non-social distancers were rated as experiencing antisocial emotions shared with other animals to a greater extent than the ingroup. However, there was no difference between climate change deniers and the ingroup for this condition. This may be because in this context, it is reasonable to suppose believers in climate change experience substantial levels of emotions such as anger and irritation. This highlights the importance of social context and stereotyping as well as ingroup preferences in explaining emotion attribution. Study 3 employed a minimal groups design in order to measure similar effects in the absence of learned stereotypes and historical negative feeling.

6. Study 3: Testing the social preference account with minimal groups

In our final experiment, we aim to replicate findings from Study 2 within a minimal groups design. Though we chose our six outgroup exemplars for Studies 1 and 2 to maximise our chances of detecting infrahumanization should it occur (based on past empirical work and also following suggestions from Leyens, 2009), it remains possible that learned stereotypes and intergroup antagonism may have weighted responses towards reflecting social preferences as opposed to subtle dehumanization. By using a minimal groups design, we were able to ensure we tested between the two competing hypotheses in the absence of these additional factors. Prior work has reported infrahumanization effects in minimal groups (Demoulin et al., 2009; Simon & Gutsell, 2020), meaning this approach was methodologically appropriate for comparing the two theories.

Participants were first allocated to novel groups using a dot estimation task (Diehl, 1990; Ratner & Amodio, 2013; Tajfel, 1970). Then, participants rated the novel ingroup and novel outgroup on emotional experiences as described for Study 2. The study design and predictions were identical as for Study 2. Infrahumanization would be observed in an interaction between group and humanness such that uniquely human emotions are more strongly ascribed to ingroup than outgroup, both for prosocial and antisocial emotions. However, we again hypothesised that prosocial human emotions would typically be attributed more strongly to the ingroup and antisocial ones to the outgroup.

6.1. Study 3 Methods

6.1.1. Participants

Based on the same power analysis as reported for Studies 1 and 2, 130 different participants completed the ratings in each experiment. Participants were eligible to

take part if they were over 18, fluent in English, and had not taken part in any of the other experiments reported presently. Nine people failed one or more attention check and their data was excluded and replaced. Of the final sample, 56 participants were female and 74 were male, aged from 18 to 57 (Mean age = 26.8, *SD* = 9.10).

6.1.2. Minimal group paradigm

Participants were assigned to novel groups based on a classic dot estimation task (e.g., Diehl, 1990; Ratner & Amodio, 2013; Tajfel, 1970). Participants were told that the purpose of the study was to help us understand the ways in which people ascribe emotions to different groups of individuals and that they would first perform a simple numerical estimation task to identify which group (out of two) they belonged to. Instead of the common categories of ‘over-estimators’ and ‘under-estimators’, we used the terms ‘spatial-estimators’ and ‘object-estimators’. This was because the emotion attribution task relied on a form of estimation (of emotional experience) and we wanted to ensure the group labels did not interact with later emotion judgments. Before the task, participants were told that people can be categorised as taking an object approach to estimation ('object-estimators') or a spatial approach to estimation ('spatial-estimators') and that individual tendencies for the two styles are equally distributed in the population.

In the dot estimation task, participants saw eleven images of random patterns of dots each on the screen for 1 second. After each image, participants had to enter the number of dots they believed they had seen before the next image appeared. The task and stimuli were based on an Inquisit script from Millisecond (https://www.millisecond.com) adapted for presentation on Qualtrics.

826 Following the task, half of the participants were told they had been classified
827 as a spatial-estimator and the other half were told they had been classified as an
828 object-estimator. This procedure met the key criteria for a minimal group paradigm
829 (Tajfel, 1970; Tajfel et al., 1971).

830 **6.1.3. Stimuli, Scales, Procedure, Design and data analysis**

831 The emotion stimuli, scale presentation, procedure, design and planned data
832 analysis were all exactly the same as outlined for Study 2.

833 **6.2. Study 3 Results**

834 **6.2.1. Blatant dehumanization and attitude scores**

835 The novel outgroup was rated as significantly less human than the novel ingroup on
836 the blatant dehumanization scale (p=.002) (Figure 1). However the novel outgroup
837 was still rated closest to the silhouette reminiscent of a modern human on the scale,
838 and was not 'blatantly dehumanized' to the extent that outgroups in Studies 1 and 2
839 were. Participants reported feeling significantly more negative towards the novel
840 outgroup than the novel ingroup on the attitude scale (p<.001). Figure S1
841 (supplementary information) shows mean results for the ingroup and outgroup on
842 the blatant dehumanization and group preference measures.

843 **6.2.2. Intergroup emotion ascription ratings**

844 For ratings towards novel ingroup and outgroup members (minimal group design),
845 there were main effects of group, $F(1, 129) = 7.58$, p =.007, η_p^2= .055, of
846 humanness, $F(1, 129) = 32.93$, p <.001, η_p^2= .203, and of sociality, $F(1, 129) =$
847 99.74, p <.001, η_p^2= .436. Ratings were higher overall for ingroup than outgroup, for
848 emotions shared with other animals than for uniquely human emotions, and for
849 prosocial than antisocial emotions.

There was a significant interaction between group and sociality, $F(1, 129) = 22.45$, $p < .001$, $\eta_p^2 = .148$, but not between group and humanness, $F(1, 129) = 1.84$, $p = .177$, $\eta_p^2 = .014$, nor between humanness and sociality, $F(1, 129) = .15$, $p = .704$, $\eta_p^2 = .001$. Pairwise comparisons showed that overall, ratings were higher for ingroup than outgroup for prosocial emotions ($p < .001$), but higher for outgroup than ingroup for antisocial emotions ($p = .007$).

All effects were qualified in a significant three-way interaction $F(1, 129) = 13.25$, $p < .001$, $\eta_p^2 = .093$. Planned comparisons showed that ratings were higher for ingroup than outgroup on prosocial terms, both uniquely human and shared with other animals ($ps < .001$). Ratings were higher for outgroup than ingroup on uniquely human antisocial terms ($p < .001$), but for antisocial terms shared with other animals there was no difference between ingroup and outgroup ($p = .637$).

Mean scores (M) and standard errors of the mean (SE) for each of the conditions in Study 3 are shown in, supplementary information, Table S4. Figure 4 shows results for Study 3.

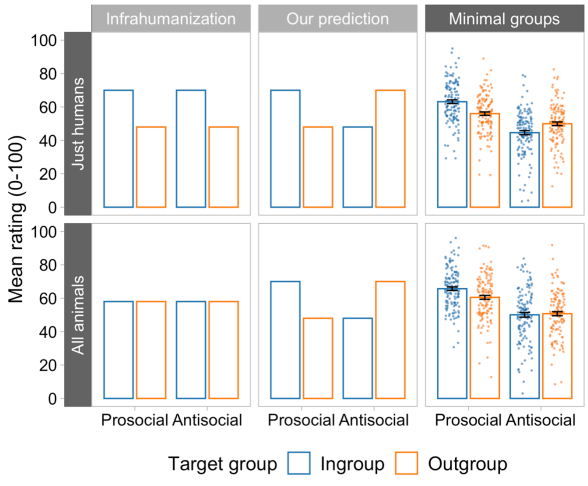

Figure 4. Evidence for social preference but not infrahumanization in a minimal group design. Contrary to infrahumanization theory and in line with the

social preference account ratings were higher for the ingroup than the outgroup for prosocial uniquely human emotions but higher for the outgroup than the ingroup for antisocial uniquely human emotions. Error bars represent standard errors.

6.3. Study 3 Discussion

Contrary to the predictions of infrahumanization theory novel outgroup members were not denied uniquely human emotions relative to novel ingroup members. There was no significant interaction between group and humanness but there was an interaction between group and sociality. Outgroup members were thought to experience prosocial uniquely human emotions to a lesser extent than ingroup members, but antisocial uniquely human emotions to a greater extent than ingroup members. This finding shows that even in a novel group context free from learned stereotypes and antagonism, our social preference account better explains intergroup biases in emotion attribution than infrahumanization theory.

7. General discussion

We found no convincing evidence for infrahumanization. In our first set of studies we broadly replicated previously reported effects (Study 1) showing our paradigm was well placed to detect infrahumanization if it occurs. Our subsequent results suggest that, in the seven intergroup contexts we employed, what appeared to be evidence for infrahumanization can be better explained by social preference (Study 2). When emotion terms varied on sociality rather than on valence, people did not 'subtly dehumanize' the outgroups we included by denying them uniquely human emotions. Rather, they attributed prosocial emotions more strongly to ingroup members and antisocial emotions more strongly to outgroup members, regardless of humanness.

892 This accords with recent critiques of the social psychological literature on
893 dehumanization more generally (Bloom, 2017; Lang, 2010, 2020; Manne, 2016,
894 2018; Over, 2020a, 2020b; Smith, 2014, 2016) and offers an important conceptual
895 development to our understanding of intergroup bias in emotion judgements.
896 In practical terms, more accurately characterising the ways in which the
897 emotions of different groups are perceived has important implications for real world
898 settings such as criminal justice, in which certain defendants, for example those
899 perceived as belonging to a religious outgroup, might be unfairly viewed as
900 possessing lower levels of remorse but also higher levels of contempt as a result of
901 their group membership. One of the main reasons why infrahumanization theory has
902 been influential in intergroup relations research is because it has been causally
903 linked to negative behavioural consequences. For example, previous research has
904 suggested that infrahumanizing outgroups reduces prosocial behaviour towards
905 them (Cuddy et al., 2007; Vaes et al., 2002, 2003). In light of the present findings,
906 future research would benefit from revisiting previously-reported links between
907 biases in emotion attribution and prosocial and antisocial behaviours.
908 Our results dovetail with recent empirical work that challenges the predictions
909 made by Haslam's (2006) dual model of dehumanization (Enock et al., 2021). This
910 research showed that when undesirable human-specific characteristics (such as
911 'corrupt' and 'selfish') are included in overall measures of humanness, there is no
912 evidence for either animalistic or mechanistic dehumanization of outgroups as
913 characterised by the dual model. Rather, desirable human qualities are more
914 strongly attributed to ingroup members and undesirable human qualities to outgroup
915 members. The present work extends these findings by further demonstrating the

916 importance of considering sociality confounds when measuring psychological
917 processes of ‘dehumanization’, this time through another highly prominent
918 framework within the field.
919 During the review process, it was put to us that because dimensions of
920 valence and sociality correlate highly in our pretest, the two constructs are
921 “indistinguishable”, thus rendering our critique obsolete. We believe this represents
922 a misunderstanding. Height and weight are strongly positively correlated, yet they
923 are distinct constructs. Similarly, even though emotions that are generally perceived
924 as prosocial may also perceived as positive to experience, and emotions that are
925 generally perceived as antisocial may also be perceived as negative to experience,
926 the two constructs are clearly conceptually distinct. While sadness is negative to
927 experience, it is not inherently antisocial in character. Schadenfreude on the other
928 hand is, by definition, positive to experience but antisocial in character. Many
929 research findings converge on the view that while ‘regret’ or ‘remorse’ are negative
930 to feel they are not unkind in character (Stearns & Parrott, 2012, see also Parkinson,
931 1996; Vaish & Hepach, 2020; van Kleef et al., 2016). Our argument is that the kinds
932 of negative emotions that proponents of infrahumanization theory have included in
933 previous research, such as disillusion, regret, melancholy and remorse (e.g., Banton
934 et al., 2020; Leyens et al., 2001; Paladino et al., 2002; Vaes et al., 2003) are
935 negative to experience but are not antisocial and so are not best placed to
936 determine whether infrahumanization is separate from intergroup preference.
937 While our results offer an important and novel empirical critique of prior work
938 on infrahumanization, we acknowledge that we only tested seven intergroup
939 contexts. Our social contexts varied in animosity, with some containing high prior

940 animosity (criminals and child molesters) and some containing little or no prior
941 animosity (minimal groups). Despite this, without testing many more groups (such
942 as those based on nationality, race or gender), it remains a possibility that
943 infrahumanization could sometimes occur even when sociality of emotion is
944 controlled. However, participants explicitly dehumanized all seven outgroups relative
945 to the ingroup on the blatant dehumanization scale, suggesting we would likely
946 observe infrahumanization if it occurs (Kteily et al., 2015). Further, the groups we
947 chose exemplify the criteria for infrahumanization proposed in prior work (Leyens,
948 2009).

949 Our results demonstrate both ingroup favouritism (assigning greater prosocial
950 feeling to the ingroup) and outgroup derogation (assigning greater antisocial feeling
951 to the outgroup) (Brewer, 1999; Hewstone et al., 2002). However, we also note that
952 group specific stereotypes and particular social contexts are likely to play an
953 important role in these processes (Fiske et al., 2002). For example, it is likely that
954 group status may affect the specific emotions that are ascribed to group members.
955 Emotions such as 'contempt' and 'schadenfreude' are included as exemplars of
956 antisocial uniquely human emptions in the present work, but it may not be the case
957 that outgroup members perceived as 'lower status' such as homeless people would
958 be attributed these antisocial emotions to a greater extent than the ingroup because
959 these particular emotions imply a position of status. There may be other uniquely
960 human yet antisocial emotions that a 'lower status' outgroup may be more likely to
961 be perceived as experiencing, such as bitterness or envy. Similarly, it is possible
962 that groups such as 'immigrants' could be perceived as feeling high levels of
963 optimism or nostalgia by nature of their situation. Our goal in this research was not

964 to explore the many nuances of emotion attribution in intergroup contexts but rather
965 more modest in scope, we aimed to show that apparent evidence for
966 infrahumanization may be better explained by other factors. Future work would
967 benefit from more closely examining the role of stereotypes and specific social
968 contexts as well as preference effects in explaining intergroup bias in emotion
969 attribution.

970 We also acknowledge that we only employed explicit measures whereas
971 infrahumanization theory has also gained support from implicit measures (Boccato
972 et al., 2007; Paladino et al., 2002). We chose to do this because explicit measures
973 have provided considerably stronger evidence for infrahumanization than have
974 implicit measures. Results from implicit measures showing stronger associations
975 between certain groups and particular emotion terms are inherently ambiguous. It is
976 not clear whether automatic associations reflect estimates of the emotions the group
977 experiences, or whether they reflect participants' own emotional reaction. For
978 example, would an implicit association between 'anger' and 'immigrants' reflect a
979 belief that immigrants feel anger, or automatic anger towards immigrants? The
980 former could support infrahumanization theory but there is no way to rule out the
981 latter. Nevertheless, the field would benefit from careful empirical research
982 rigorously controlling for emotion sociality in more implicit contexts. Additionally, we
983 acknowledge that, following the majority of prior work on infrahumanization, we
984 conceptualised humanness and valence/sociality dichotomously and as such tested
985 our predictions with relatively few exemplars from each category. Though this
986 approach is standard in the field, future work may benefit from testing similar

hypotheses but treating humanness and sociality as continuous predictors (see Castano & Giner-Sorolla, 2006).

These possibilities do not detract from the central importance of our critique. More generally, our results illustrate the importance of considering the sociality of emotion terms employed as stimuli, a methodological advancement that will be crucial to incorporate in any future studies of emotion attribution in other intergroup contexts. To accurately test for 'infrahumanization', future research must consider the central role of emotion sociality as separate from emotional experience. Prior work has reported infrahumanization to be extremely widespread in society and prevalent across a multitude of intergroup divides (Banton et al., 2020; Cortes et al., 2005; Cuddy et al., 2007; Gaunt, 2009; Leyens et al., 2000, 2001; Rodríguez-Pérez et al., 2011; Simon & Gutsell, 2020; Vaes et al., 2002, 2003). Rigorous measurement, tighter experimental control and more careful consideration of social context may change some or all of the conclusions from previous research.

If psychological research is to effectively inform intervention to improve intergroup relations, it is essential it accurately characterises the underlying mechanisms of intergroup bias. Our findings suggest the construct of infrahumanization may obscure more than it reveals about intergroup bias.

References

Appiah, K. A. (2008). *Experiments in ethics*. Harvard University Press.

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). *The Root of All Cruelty?* The New Yorker. https://www.newyorker.com/magazine/2017/11/27/the-root-of-all-cruelty

Boccato, G., Cortes, B. P., Demoulin, S., & Leyens, J.-P. (2007). The automaticity of infra-humanization. *European Journal of Social Psychology*, *37*(5), 987–999.

Brown, R., Eller, A., Leeds, S., & Stace, K. (2007). Intergroup contact and intergroup attitudes: A longitudinal study. *European Journal of Social Psychology*, *37*(4), 692–703.

Calfano, B. (2018). *Muslims, Identity, and American Politics*. Routledge.

Capozza, D., Trifiletti, E., Vezzali, L., & Favara, I. (2013). Can intergroup contact improve humanity attributions? *International Journal of Psychology*, *48*(4), 527–541.

Castano, E., & Giner-Sorolla, R. (2006). Not quite human: Infrahumanization in response to collective responsibility for intergroup killing. *Journal of Personality and Social Psychology*, *90*(5), 804.

Cortes, B. P., Demoulin, S., Rodriguez, R. T., Rodriguez, A. P., & Leyens, J.-P. (2005). Infrahumanization or Familiarity? Attribution of Uniquely Human

Emotions to the Self, the Ingroup, and the Outgroup. *Personality and Social Psychology Bulletin*, *31*(2), 243–253. https://doi.org/10.1177/0146167204271421

Cuddy, A. J. C., Rock, M. S., & Norton, M. I. (2007). Aid in the Aftermath of Hurricane Katrina: Inferences of Secondary Emotions and Intergroup Helping. *Group Processes & Intergroup Relations*, *10*(1), 107–118. https://doi.org/10.1177/1368430207071344

Demoulin, S., Cortes, B., Viki, T., Rodríguez, A., Rodriguez, R., Paladino, M., & Leyens, J.-P. (2009). The Role of In-group Identification in Infra-humanization. *International Journal of Psychology : Journal International de Psychologie*, *44*, 4–11. https://doi.org/10.1080/00207590802057654

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

Diehl, M. (1990). The minimal group paradigm: Theoretical explanations and empirical findings. *European Review of Social Psychology*, *1*(1), 263–292.

Ekman, P. (1992). An argument for basic emotions. *Cognition and Emotion*, *6*(3–4), 169–200. https://doi.org/10.1080/02699939208411068

Enock, F. E., Flavell, J. C., Tipper, S. P., & Over, H. (2021). No convincing evidence outgroups are denied uniquely human characteristics: Distinguishing intergroup preference from trait-based dehumanization. *Cognition*, *212*, 104682.

Gaunt, R. (2009). Superordinate Categorization as a Moderator of Mutual Infrahumanization. *Group Processes & Intergroup Relations*, *12*(6), 731–746. https://doi.org/10.1177/1368430209343297

Goff, P. A., Eberhardt, J. L., Williams, M. J., & Jackson, M. C. (2008). Not yet human: Implicit knowledge, historical dehumanization, and contemporary consequences. *Journal of Personality and Social Psychology*, *94*(2), 292–306. https://doi.org/10.1037/0022-3514.94.2.292

Goff, P. A., Jackson, M. C., Di Leone, B. A. L., Culotta, C. M., & DiTomasso, N. A. (2014). The essence of innocence: Consequences of dehumanizing Black children. *Journal of Personality and Social Psychology*, *106*(4), 526.

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., & 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., & 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., & Schmid, K. (2014). Neighbourhood ethnic diversity and orientations toward Muslims in Britain: The role of intergroup contact. *The Political Quarterly*, *85*(3), 320–325.

Johnson, N. F., Velásquez, N., Restrepo, N. J., Leahy, R., Gabriel, N., El Oud, S., Zheng, M., Manrique, P., Wuchty, S., & Lupu, Y. (2020). The online competition between pro- and anti-vaccination views. *Nature*, *582*(7811), 230–233. https://doi.org/10.1038/s41586-020-2281-1

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

Kteily, N., Hodson, G., & Bruneau, E. (2016). They see us as less than human: Metadehumanization predicts intergroup conflict via reciprocal dehumanization. *Journal of Personality and Social Psychology*, *110*(3), 343.

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*.

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

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

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.

Over, H. (2020a). Falsifying the Dehumanization Hypothesis. *Perspectives on Psychological Science*, 1745691620969657.

Over, H. (2020b). Seven challenges for the dehumanisation hypothesis. *Perspectives on Psychological Science*.

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

Parkinson, B. (1996). Emotions are social. *British Journal of Psychology*, *87*(4), 663–683. https://doi.org/10.1111/j.2044-8295.1996.tb02615.x

Prati, F., Crisp, R. J., Meleady, R., & Rubini, M. (2016). Humanizing outgroups through multiple categorization: The roles of individuation and threat. *Personality and Social Psychology Bulletin*, *42*(4), 526–539.

Prosser, A. M. B., Judge, M., Bolderdijk, J. W., Blackwood, L., & Kurz, T. (2020). 'Distancers' and 'non-distancers'? The potential social psychological impact of moralizing COVID-19 mitigating practices on sustained behaviour change. *British Journal of Social Psychology*, *59*(3), 653–662. https://doi.org/10.1111/bjso.12399

Rai, T. S., Valdesolo, P., & Graham, J. (2017). Dehumanization increases instrumental violence, but not moral violence. *Proceedings of the National Academy of Sciences*, *114*(32), 8511–8516.

Ratner, K. G., & Amodio, D. M. (2013). Seeing "us vs. them": Minimal group effects on the neural encoding of faces. *Journal of Experimental Social Psychology*, *49*(2), 298–301. https://doi.org/10.1016/j.jesp.2012.10.017

Rodríguez-Pérez, A., Delgado-Rodríguez, N., Betancor-Rodríguez, V., Leyens, J.-P., & Vaes, J. (2011). Infra-humanization of outgroups throughout the world. The role of similarity, intergroup friendship, knowledge of the outgroup, and status. *Anales de Psicología*, *27*(3), 679–687.

Simon, J. C., & Gutsell, J. N. (2020). Effects of minimal grouping on implicit prejudice, infrahumanization, and neural processing despite orthogonal social categorizations. *Group Processes & Intergroup Relations*, *23*(3), 323–343.

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, D. L. (2020). *On inhumanity: Dehumanization and how to resist it*. Oxford University Press.

Stearns, D. C., & Parrott, W. G. (2012). When feeling bad makes you look good: Guilt, shame, and person perception. *Cognition and Emotion*, *26*(3), 407–430. https://doi.org/10.1080/02699931.2012.675879

Tajfel, H. (1970). Experiments in intergroup discrimination. *Scientific American*, *223*(5), 96–103.

Tajfel, H., Billig, M. G., Bundy, R. P., & Flament, C. (1971). Social categorization and intergroup behaviour. *European Journal of Social Psychology*, *1*(2), 149–178. https://doi.org/10.1002/ejsp.2420010202

Tam, T., Hewstone, M., Cairns, E., Tausch, N., Maio, G., & Kenworthy, J. (2007). The impact of intergroup emotions on forgiveness in Northern Ireland. *Group Processes & Intergroup Relations*, *10*(1), 119–136.

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 dehumanisation 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

Supplementary Information for:

Intergroup preference, not dehumanization, explains social biases in emotion attribution

Authors: Florence E. Enock[1], Steven P. Tipper[1], and Harriet Over[1]

[1]Department of Psychology, University of York, York, United Kingdom. YO10 5DD.

*Correspondence to: Florence Enock, Florence.Enock@york.ac.uk.

Supplementary information for pretest

Supplementary information on pretest scales

To determine the extent to which emotions are considered to be uniquely human or shared with other species, positive or negative to experience, and prosocial or antisocial in character, participants rated fifty-four emotion words on three scales.

The words we included were: anger, disgust, fear, happiness, sadness, surprise (these first six are often considered primary emotions), admiration, arrogance, bitterness, compassion, complacency, conceit, contempt, contentment, disappointment, disillusion, embarrassment, empathy, envy, friendliness, gloating, greed, grief, guilt, hatred, hope, hopelessness, hostility, humiliation, humility, irritation, jealousy, loneliness, love, melancholy, nervousness, nostalgia, optimism, patience, pride, regret, relief, remorse, resentment, resignation, schadenfreude, scorn, self-satisfaction, shame, shyness, smugness, spite, tenderness, vengefulness. These terms were obtained from prior work on infrahumanization and from emotion research more generally (Demoulin et al., 2004; Leyens et al., 2001).

The Humanness scale asked: “Using the slider, please indicate how much the emotion in each of the following questions is experienced by humans compared to other species (i.e., is this emotion unique to humans?)” The bottom end of the slider, 0, corresponded to *Just other species* and the top end, 100, corresponded to *Just humans*, with the midpoint, 50, indicating *Equal to humans and other species.*

The Valence scale asked: “Using the slider, please indicate what you think this emotion is like to experience (i.e., how does it make you feel?).” The bottom end of the slider, 0, corresponded to *Extremely negative* and the top end, 100,

corresponded to *Extremely positive*, with the midpoint, 50, indicating *Neutral to experience*.

The Sociality scale asked: “Using the slider, please indicate what you think someone who regularly experiences this emotion is like (i.e., how kind are they likely to be?).” The bottom end of the slider, 0, corresponded to *Extremely unkind* and the top end, 100, corresponded to *Extremely kind*, with the midpoint, 50, indicating *Neither kind nor unkind*.

Taking our lead from infrahumanization theory, we were interested in lay conceptions of emotions. As prosociality is not a common word for the general population, we use ‘kindness’ in our scale to capture ‘what you think someone who regularly experiences this emotion is like’ as opposed to ‘what you think this emotion is like to experience’. We use the term ‘sociality’ throughout to clearly distinguish from ‘valence’ as subjective experience.

Each item was scored from 0-100 but participants could not see the numbers. The three scales were presented in separate blocks on sequential screens and the order of completion was counterbalanced such that one third of participants rated Humanness then Valence then Sociality, one third rated Sociality then Humanness then Valence, and one third rated Valence then Sociality then Humanness. The emotion items within each block were randomised. One attention check per block was included approximately halfway through, such as ‘Please indicate extremely positive’. Participants were excluded and their data replaced if they failed one or more attention checks.

Supplementary results for pretest

The mean ratings for each emotion on Humanness, Valence and Sociality are presented in Table S1. We show ratings from most to least uniquely human, most to least positive to experience and most to least prosocial. On the humanness scale, we were most interested in finding emotions perceived as shared with other species (scoring close to 50) and those perceived as being only experienced by humans (scoring close to 100). While some emotions were rated similarly on sociality and valence (compassion was rated as highly positive to experience and highly prosocial), others were rated orthogonally (grief was rated as highly negative to experience but neither prosocial nor antisocial).

Table S1. Pretest results

Humanness			**Valence**			**Sociality**		
Most to least human			**Most to least positive**			**Most to least prosocial**		
Emotion	**M**	**SE**	**Emotion**	**M**	**SE**	**Emotion**	**M**	**SE**
Nostalgia	85.1	2.43	Happiness	95.5	1.04	Compassion	86.4	1.95
Arrogance	79.9	2.61	Love	91.7	1.72	Empathy	86.3	1.71
Optimism	79.8	2.70	Friendliness	85.0	2.12	Love	84.6	2.06
Schadenfreude	79.3	2.86	Optimism	84.6	1.67	Friendliness	82.9	2.48
Disillusion	79.2	2.39	Hope	81.0	2.05	Happiness	80.1	2.34
Humility	78.6	2.46	Compassion	80.9	2.04	Patience	75.5	2.10
Contempt	76.9	2.48	Empathy	79.1	2.09	Optimism	74.7	2.16
Regret	76.7	2.71	Relief	78.7	2.57	Tenderness	74.7	2.74
Melancholy	76.5	2.69	Self-satisfaction	78.6	2.68	Hope	71.1	1.92
Scorn	76.1	2.74	Tenderness	77.0	2.47	Admiration	70.4	2.08
Smugness	75.9	2.55	Admiration	76.9	1.86	Contentment	68.1	2.24
Humiliation	75.8	2.80	Patience	72.9	2.25	Humility	67.5	2.90
Remorse	75.6	2.41	Contentment	72.8	3.20	Relief	66.0	1.98
Embarrassment	75.1	2.53	Pride	69.8	2.69	Nostalgia	64.6	2.07
Hope	74.6	2.80	Surprise	65.4	1.86	Shyness	59.4	1.90
Greed	74.6	2.68	Humility	62.8	3.63	Surprise	58.6	1.50
Hopelessness	74.3	2.76	Nostalgia	60.8	2.93	Self-satisfaction	58.3	2.60
Conceit	74.3	2.65	Complacency	55.1	3.15	Complacency	53.4	2.51
Bitterness	74.1	2.31	Gloating	44.9	3.43	Guilt	53.0	2.34
Vengefulness	72.3	2.74	Smugness	43.6	3.55	Remorse	52.5	2.90
Resentment	71.5	2.82	Shyness	40.4	1.82	Regret	52.5	2.38
Gloating	71.3	2.60	Conceit	36.2	2.86	Pride	51.1	2.56

Hate	71.1	2.62	Schadenfreude	35.6	3.14	Nervousness	49.5	1.87
Guilt	70.5	2.83	Melancholy	34.8	2.63	Embarrassment	49.0	1.89
Complacency	70.4	2.64	Remorse	33.5	2.93	Shame	48.3	2.28
Spite	70.3	2.87	Nervousness	31.8	2.00	Grief	48.1	2.44
Shame	70.2	2.76	Contempt	31.1	2.99	Melancholy	47.6	2.30
Self-satisfaction	70.1	2.66	Resignation	26.8	2.13	Sadness	46.4	2.49
Disappointment	69.1	2.27	Disillusion	25.5	2.43	Fear	46.2	1.71
Envy	68.6	2.65	Guilt	24.9	2.12	Loneliness	45.6	2.29
Pride	68.5	2.69	Irritation	24.3	1.76	Resignation	43.8	1.81
Disgust	67.9	2.84	Arrogance	24.2	2.79	Hopelessness	42.4	2.26
Resignation	67.2	2.59	Embarrassment	23.4	2.10	Disillusion	40.7	2.14
Admiration	66.7	2.57	Regret	23.4	2.12	Disappointment	40.3	2.44
Shyness	66.4	2.40	Fear	23.3	2.28	Gloating	39.0	3.54
Relief	65.0	2.49	Spite	23.3	2.69	Smugness	35.4	3.13
Compassion	64.9	2.40	Envy	23.0	1.94	Contempt	33.8	2.65
Jealousy	64.7	2.50	Scorn	22.1	2.22	Humiliation	33.7	2.80
Patience	64.5	2.48	Resentment	21.9	1.70	Conceit	32.6	2.66
Empathy	63.4	2.87	Greed	21.6	2.39	Resentment	31.1	2.06
Contentment	62.6	2.25	Vengefulness	21.3	2.42	Irritation	29.1	2.16
Nervous	60.2	2.32	Hostility	21.3	2.89	Bitterness	28.2	2.00
Friendliness	58.6	2.12	Shame	20.8	1.85	Schadenfreude	28.1	2.89
Grief	58.5	1.67	Jealousy	19.7	1.98	Envy	27.1	2.04
Love	58.4	1.95	Disappointment	18.9	1.93	Jealousy	25.6	2.12
Anger	57.2	2.31	Grief	18.7	2.38	Scorn	25.6	2.43
Irritation	57.1	2.34	Sadness	18.1	2.42	Disgust	24.5	2.01
Surprise	56.8	1.76	Bitterness	17.9	1.62	Greed	23.9	2.26
Tenderness	56.1	1.63	Anger	17.8	1.87	Spite	22.9	2.29
Happiness	54.7	1.66	Loneliness	15.9	2.13	Anger	22.7	2.36
Sadness	54.3	1.53	Disgust	15.5	1.80	Hostility	21.6	2.96
Loneliness	53.8	1.71	Humiliation	12.9	1.75	Arrogance	20.9	1.98
Hostility	53.6	2.19	Hopelessness	12.5	1.81	Vengefulness	14.2	1.88
Fear	48.6	1.58	Hate	7.4	1.30	Hate	10.0	1.70

Table S1. Emotion terms scored from highest to lowest along dimensions of Humanness, Valence, and Sociality. Mean scores (M) and standard error of the mean (SE) are presented alongside each word. Respective to each scale, 100 indicated the emotion was highly unique to humans / extremely positive to experience / extremely kind (prosocial). 0 indicated the emotion was unique to other species / extremely negative to experience / extremely unkind (antisocial). 50 indicated the emotion applied equally to humans and other species / was neither positive nor negative to experience / neither kind nor unkind.

Supplementary information on intergroup contexts (Studies 1 and 2)

We employed six intergroup contexts across the experiments in Studies 1 and 2. In Experiments 1a and 2a, the outgroup was Muslims and the ingroup was Christians. In Experiments 1b and 2b, the outgroup was 'Individuals with criminal convictions' (criminals) and the ingroup was 'Individuals with no criminal history'. In Experiments 1c and 2c, the outgroup was 'Child molesters' and the ingroup was 'Individuals with no criminal history'. In Experiments 1d and 2d, the outgroup was 'Individuals who are against vaccination ('anti-vaxxers')' and the ingroup was 'Individuals who are in favour of vaccination ('pro-vaxxers')'. In Experiments 1e and 2e, the outgroup was 'Individuals who do not adhere to the government regulations on social distancing/quarantine during the COVID-19 pandemic' ('non-social distancers') and ingroup was 'Individuals who do adhere to the government regulations on social distancing/quarantine during the COVID-19 pandemic' ('social distancers'). In Experiments 1f and 2f, the outgroup was 'Individuals who do not believe in climate change ('climate change deniers')' and the ingroup was 'Individuals who believe in climate change ('climate change believers')'.

Supplementary information for Study 1

Supplementary information on stimuli development (Study 1)

We chose emotions from our pretest data that best fit the four emotion categories of interest. We chose four rated as highly positive and four rated as highly negative from both the most and least uniquely human terms.

For conceptual consistency between our stimulus set and the original work (Leyens et al., 2001), we omitted negative emotions that were also rated as

antisocial. This was because we suggest that the kinds of negative emotions that have been included in previous infrahumanization research (e.g., regret, melancholy, disillusion, remorse - Banton et al., 2020; Leyens et al., 2001; Paladino et al., 2002; Vaes et al., 2003) are negative to experience but are not antisocial. To first replicate infrahumanization effects, we included uniquely human emotions that were similar to ones used in prior work.

In developing the items for our emotion categories, we ensured that humanness ratings did not significantly differ between the positive and negative conditions for each level of humanness so that we could accurately separate valence effects from humanness. For example, whilst Grief was rated as less unique to humans and also more negative to experience than Nervousness, when we included Grief in the stimuli, the set of non-uniquely human positive emotions was overall higher in humanness than the non-uniquely human negative emotions. We included Nervousness instead so as to ensure the non-uniquely human positive and negative emotions were matched on perceived Humanness.

In support of our experimental manipulations, paired t-tests showed that combined, the uniquely human emotion words were rated as significantly more human (M = 78.3 ± 1.71) than emotions shared with other species (M = 55.3 ± .96), $t(59) = 14.70$, $p < .001$, $d = 1.90$. Additionally, The positive words (M = 77.3 ± 1.09) were rated as significantly more positive than the negative words (M = 25.8 ± 1.37), $t(59) = 28.89$, $p < .001$, $d = 3.73$. Humanness scores were comparable for positive (M = 79.5 ± 1.80) and negative words (M = 77.0 ± 1.90) unique to humans, $t(59) = 1.75$, $p = .085$, $d = 0.23$ and for positive (M = 56.5 ± 1.15) and negative (M = 54.2 ± 1.33) words shared with other species, $t(59) = 1.42$, $p = .159$, $d = 0.18$.

Supplementary information on design and data analysis (Study 1)

For each experiment, there were eight conditions in total in a 2 x 2 x 2 within-subjects design. Our overall design mirrored original work (Leyens et al., 2001) with some minor methodological developments. We measured emotion attribution by asking participants to indicate on a sliding scale the extent to which they believed each group to experience the emotion items (from *not at all* to *very strongly*) rather than by asking them to simply choose whether or not particular emotions applied to ingroups or outgroups. This was to provide potential for greater distribution in responses so that data were likely to be more appropriate for parametric statistics than in original studies, where ANOVAs were performed on counts from 0-3 in each condition.

We employed a within-subjects design, only including one side of each intergroup context, omitting 'group membership' as an additional between-subjects factor. Previous work found infrahumanization on both sides of group memberships and showed effects do not rest on group status (Rodríguez-Pérez et al., 2011). Further, effects have been detected in within- as well as between-subjects designs (Cortes et al., 2005). This gave us greater statistical power and the opportunity to test intergroup contexts in which it may be more difficult to obtain data from both sides.

Supplementary information for Study 2

Supplementary information on stimuli development (Study 2)

Similar to Study 1, we chose emotions from our pretest data that best fit the four emotion categories of interest. We chose four rated as prosocial and four rated as

antisocial from both the most and least uniquely human terms. Though disgust was rated as somewhat more unique to humans than the other emotions categorised as shared with other species, it was included in the set of antisocial emotions shared with other species because it is widely considered a primary (or 'basic') emotion and also because with a mean score of (M = 67.9 ± 2.84), it still fell closer to the 'equal to humans and other species' than the 'just humans' mark. While arrogance, friendliness and humility may be considered as either traits or emotions by some accounts, they are included because prior work on infrahumanization often includes trait terms as well as more traditional emotions (Capozza et al., 2013; Hodson & Costello, 2007; Vaes & Paladino, 2010). Thus, infrahumanization effects have previously been understood and reported across emotions, traits, and even simple category words such as 'wife' and 'pet' (Viki et al., 2006).

Similar to Study 1, we ensured that humanness ratings did not significantly differ between the prosocial and antisocial conditions for each level of humanness so that we could accurately separate sociality effects from ones of humanness.

Paired t-tests showed that combined, the uniquely human words were rated as significantly more human (M = 78.8 ± 1.72) than the words shared with other species (M = 57.8 ± 1.07), $t(59) = 13.87$, $p < .001$, $d = 1.79$, and the prosocial words (M = 75.00 ± 1.31) were rated as significantly more prosocial than the antisocial words (M = 25.8 ± 1.25), $t(59) = 21.61$, $p < .001$, $d = 2.79$. Humanness scores were comparable for prosocial (M = 79.5 ± 1.80) and antisocial (M = 78.0 ± 2.18) words unique to humans, $t(59) = .73$, $p = .470$, $d = 0.09$ and for prosocial (M = 56.9 ± 1.26) and antisocial (M = 58.9 ± 1.60) words shared with other species, $t(59) = 1.04$, $p = .302$, $d = 0.13$.

A paired-samples t-test showed no significant difference in valence between the negative uniquely human emotions in Study 1 (mean valence = 29.28, SE = 1.63) and the antisocial uniquely human emotions in Study 2 (mean valence = 28.23, SE = 1.74), t(59) = .516, p = .607, d = .07, with a Bayes factor of 6.23 supporting the null. However, the antisocial uniquely human emotions in Study 2 (mean sociality = 27.09, SE = 1.50) were significantly more antisocial than the negative uniquely human emotions in Study 1 (mean sociality = 48.33, SE = 1.59), t(59) = .8.63, p < .001, d = 1.11. Thus, the meaningful difference between the uniquely human terms included in Studies 1 and 2 was in the sociality – a factor not considered in previous infrahumanization research.

Supplementary Figure S1

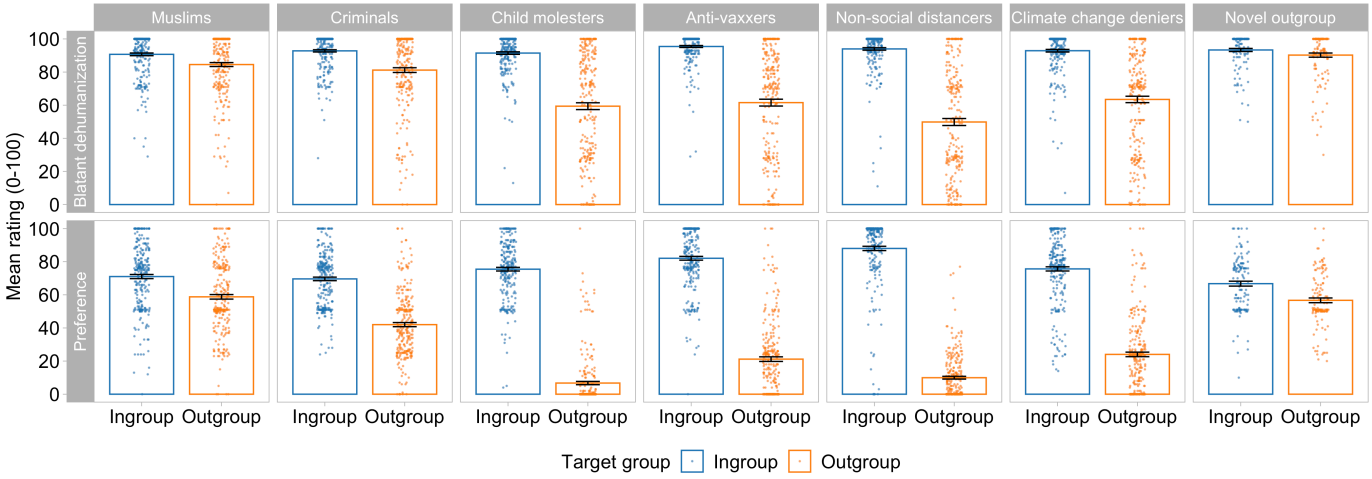

Figure S1. Blatant dehumanization and preference ratings across all experiments in Studies 1, 2 and 3. Outgroups are: Muslims (1a&2a), criminals (1b&2b), child molesters (1c&2c), anti-vaxxers (1d&2d), non-social distancers (1e&2e), climate change deniers (1f&2f), and minimal outgroups (3). We collapse data across Studies 1 and 2 for the first 6 groups, giving a total N of 260 per group context for these experiments. Outgroups were rated significantly lower than ingroups on the preference scale in all experiments (all *ps* <.001) and also as significantly 'less human' than ingroups

on the blatant dehumanization scale (all *ps* <.001 in Sudies 1 and 2, *p*=.002 in Study 3). Error bars represent standard errors.

Supplementary Tables - Studies 1, 2 and 3

Table S2. Mean emotion attribution scores by condition in Study 1

	Emotions unique to humans				Emotions shared with other animals			
	Positive	Positive	Negative	Negative	Positive	Positive	Negative	Negative
Expt.	Ingroup	Outgroup	Ingroup	Outgroup	Ingroup	Outgroup	Ingroup	Outgroup
1a	72.6 (1.08)	63.9 (1.34)	56.5 (1.55)	52.1 (1.53)	71.2 (1.11)	62.7 (1.44)	55.4 (1.63)	57.8 (1.59)
1b	67.2 (1.16)	50.8 (1.28)	57.2 (1.37)	62.3 (1.42)	72.1 (1.26)	50.1 (1.59)	62.1 (1.56)	66.8 (1.40)
1c	67.9 (1.16)	35.3 (1.36)	56.2 (1.28)	44.1 (1.62)	70.4 (1.19)	37.9 (1.74)	60.2 (1.46)	57.7 (1.78)
1d	67.0 (1.20)	51.4 (1.61)	43.5 (1.83)	52.7 (1.70)	67.4 (1.18)	56.7 (1.50)	48.1 (1.72)	60.9 (1.66)
1e	62.0 (1.34)	40.4 (1.37)	47.0 (1.42)	35.3 (1.37)	53.5 (1.44)	40.7 (1.24)	73.2 (1.46)	35.1 (1.67)
1f	58.5 (1.11)	50.5 (1.26)	60.7 (1.16)	41.3 (1.45)	58.0 (1.22)	52.7 (1.27)	65.9 (1.24)	48.1 (1.72)

Standard errors of the mean are in parentheses.

Table S3. Mean emotion attribution scores by condition in Study 2

	Emotions unique to humans				Emotions shared with other animals			
	Prosocial	Prosocial	Antisocial	Antisocial	Prosocial	Prosocial	Antisocial	Antisocial
Expt.	Ingroup	Outgroup	Ingroup	Outgroup	Ingroup	Outgroup	Ingroup	Outgroup
2a	71.9 (1.07)	66.0 (1.22)	45.6 (1.50)	53.2 (1.39)	75.8 (1.16)	67.0 (1.47)	48.9 (1.55)	57.8 (1.69)
2b	69.4 (91.08)	50.6 (1.34)	52.1 (1.22)	62.0 (1.15)	72.3 (1.20)	48.6 (1.51)	54.9 (1.31)	71.7 (1.07)
2c	64.7 (1.14)	37.3 (1.64)	51.4 (1.30)	61.8 (1.43)	68.9 (1.36)	35.1 (1.74)	56.2 (1.46)	61.4 (1.27)
2d	66.1 (1.18)	51.2 (1.50)	47.2 (1.52)	64.2 (1.43)	73.5 (1.26)	52.9 (1.73)	51.3 (1.83)	74.7 (1.35)
2e	62.6 (1.49)	39.8 (1.24)	32.3 (1.37)	63.5 (1.41)	60.4 (1.66)	36.4 (1.31)	48.1 (1.53)	61.7 (1.39)
2f	58.4 (1.36)	51.2 (1.22)	46.4 (1.45)	63.4 (1.21)	67.2 (1.35)	51.3 (1.48)	64.7 (1.40)	62.2 (1.46)

Standard errors of the mean are in parentheses.

Table S4. Mean emotion attribution scores by condition in Study 3

	Uniquely human emotions		Emotions shared with other animals	
	Ingroup	Outgroup	Ingroup	Outgroup
Prosocial	63.1 (1.09)	56.0 (1.02)	65.7 (1.07)	60.5 (1.18)
Antisocial	44.6 (1.20)	50.0 (1.09)	50.1 (1.39)	50.7 (1.21)

Standard errors of the mean are in parentheses.