Dehumanization after all: Distinguishing intergroup evaluations from trait-based dehumanization

Jeroen Vaes

University of Trento, Italy

Corresponding author:

Jeroen Vaes, University of Trento, Department of Psychology and Cognitive Sciences, Corso Bettini , 38068 Rovereto, Italy Email: jeroen.vaes@unitn.it Tel. +39 0464808639

Word count (incl. abstract, excluding figures and references): 3493

Abstract

Over the last two decades, a plethora of research using a large variety of measures and paradigms has demonstrated how people often tend to dehumanize members of certain outgroups. When trait-based measures were used, ingroup members showed to attribute less human positive and negative characteristics to members of the outgroup. Enock and colleagues (2021) recently criticized these findings stating that previous work on this topic has investigated the attribution of mostly socially desirable human traits making it impossible to determine that trait-based dehumanization is distinct from intergroup preferences. In a set of studies in which they balanced the human and non-human traits for valence, they did not find any evidence for outgroup dehumanization. Using the same experimental material, the current work conceptually replicated a subsample of these studies in two pre-registered experiments introducing a more parsimonious measure of dehumanization that is based on a large number of traits and takes both the variance in typicality and humanness judgments into account, at the same time allowing for a rigid control of the valence of the traits. Results clearly indicated the presence of a dehumanization effect in both studies over and above intergroup evaluations. As such, these results are in line with previous work on dehumanization and highlight the risks of gauging dehumanization through the attribution of a small number of fixed human traits that are not controlled for their desirability.

1. Introduction

Over the last two decades a consistent body of research has shown how people often tend to dehumanize members of certain outgroups (see Bain, Vaes, & Leyens, 2014; Haslam & Loughnan, 2014; Kteily & Landry, 2022; Vaes et al., 2012, for reviews). The common reported finding in this research is that ingroup members attribute less human positive and negative characteristics to members of the outgroup. Over the years, this hypothesis has been corroborated using emotion-based (e.g., Leyens et al., 2001), trait-based (e.g., Bain et al., 2009), and metaphor-based measures (e.g., Goff et al., 2008; Kteilly et al., 2015), as well as perceptual (e.g., Fincher & Tetlock, 2016) and neural evidence (e.g., Harris & Fiske, 2006; Ruzzante & Vaes, 2021). Within this framework and drawing on the initial work of Leyens and colleagues (2000, 2007), Haslam (2006) proposed a dual-process model of dehumanization in which others are either described with less uniquely human characteristics (HU) likening them more closely to animals, or attributed less human nature (HN) traits characterizing them more like automata or robots. As a result, groups are animalistically or mechanistically dehumanized based on the specific kind of traits they are denied.

Recently, this perspective has been criticized empirically by Enock and colleagues (2021). These authors claimed that many of the HU and HN traits that were used in previous research were also socially desirable, such as warmth, rationality and civility. At the same time, several antisocial traits such as jealousy, arrogance, and selfishness, are clearly human traits that are probably more likely attributed to outgroup instead of ingroup members. According to them, previous research that measured dehumanization through the attribution of socially desirable traits is not fit to distinguish dehumanization biases from intergroup preferences. Based on this critique they tested an alternative hypothesis stating

that ingroup and outgroup members possess human characteristics to a similar extent, but these attributes are more negative and antisocial in character when attributed to the outgroup. In seven experiments, testing three different intergroup contexts and balancing the human and non-human traits for valence, Enock and colleagues (2021) found evidence for their alternative hypothesis. Desirable traits were ascribed more strongly to ingroup compared to outgroup members, while the reverse effect was reported for undesirable traits. In none of these studies, any significant animalistic or mechanistic dehumanization effect of the outgroup occurred.

Even though this critique and the reported results seem problematic for trait-based dehumanization research in general and the dual process model of dehumanization in particular, the paper of Enock and colleagues (2021) presents several caveats and empirical limitations. First of all, a broader look at the literature on dehumanization suggests that the majority of research on this topic controlled for valence either using a balanced sample of both positive and negative human and non-human traits or by directly controlling participants' humanness judgments for valence (e.g., Bain et al., 2009; Martinez et al., 2012; Miranda et al., 2014; Paladino & Vaes, 2009; Saminaden et al., 2010; Vaes et al., 2010; Vaes & Paladino, 2010). As such, the number of studies that has demonstrated outgroup dehumanization over and above intergroup prejudice and outgroup derogation are much higher than Enock and colleagues (2021) would like to admit. Secondly, some research has shown that if we attribute negative traits to ourselves or our ingroups, we do this especially when these traits are perceived as human (Koval et al., 2012). Indeed, admitting to our faults or shortcomings is easier to justify, the more these characteristics are perceived to be human. Finally, researchers have acknowledged the limits of gauging dehumanization through the attribution of a small number of fixed human traits that are pre-tested in a

separate sample. Even though this approach has been used in many dehumanization studies, it has some important limits for at least two reasons. First, with a small number of pre-tested traits, one can comment that the presence or absence of one or another trait does not fully represent one or the other dimension of humanness. Therefore, the use of a large number of traits that vary on the human dimension is always recommendable. Secondly, people might differ in the extent to which they perceive traits to be human and these judgments might vary as a function of the context. Specifically, Paladino and Vaes (2009) manipulated whether traits and emotions were typical of the ingroup or the outgroup and measured the perceived humanness of these traits. In three different intergroup contexts, they found that the traits that were described typical of the ingroup were judged to be more human compared to when the same traits were said to describe the outgroup. This variability, however, is not taken into account when traits are independently selected as high or low in humanness on the basis of a separate pre-test. To overcome this shortcoming Vaes and Paladino (2010), measured how typical a large set of traits was perceived to describe a specific group, while at the same time measuring both the humanness and valence of these traits in nine different intergroup situations. Calculating the within-participant correlations between typicality and humanity judgments from which they partialled out valence, they managed to obtain a dehumanization index that is based on a large number of traits, that takes both the variance in typicality and humanness judgments into account, at the same time allowing for a rigid control of the valence of the traits. This and similar approaches have been successfully used in several dehumanization studies (e.g., Castano & Giner-Sorolla, 2006; Deluca-McLean & Castano, 2009; Miranda et al., 2014; Morris et al., 2018; Vaes et al., 2010).

2. A two-study preregistered conceptual replication of Study 2a and 2b of Enock and colleagues (2021)

Given the shortcomings identified in Enock and colleagues' (2021) work, the current studies aimed to conceptually replicate some of their studies using the more parsimonious approach introduced by Vaes and Paladino (2010). Apart from the specific methodology, we aimed to keep the experimental material as similar as possible. Therefore, we used the full set of 60 traits that were pretested by Enock and colleagues (2021) that contained the negative human traits that according to them were missing in previous work on dehumanization. We asked to judge these traits on the same dimensions (typicality in ascribing these traits to the ingroup and the outgroup, humanness (HU in Study 1a and HN in Study 1b) and valence. To judge these dimensions, we asked participants the same questions using the same wording that was originally used in the research of Enock and colleagues (2021). Finally, we used the same participant pool and outgroup (i.e., immigrants) as they used in their Study 2a and 2b. It is important to note, that participants were not selected on the basis of their political affiliation as Enock and colleagues (2021) did. They decided to only include right-wing participants. Even though political affiliation has shown to moderate people's tendency to dehumanize immigrants in previous research (Markowitz & Slovic, 2020), the dehumanization of immigrants has been observed controlling for political attitudes (Bruneau et al., 2018) and can therefore not be considered a necessary precursor to observe the effect. Both studies were pre-registered and the full pre-registration can be found here (https://osf.io/gwjcv/). Unlike Enock and colleagues (2021), we expected to find a general dehumanization effect showing that ingroup members were attributed more HU (Study 1a) and HN (Study 1b) traits even when controlling for their valence, compared to immigrant outgroup members.

2.1 Method

2.1.1. Participants

An a-priori power analysis indicated that to test the main hypothesis comparing the ingroup and the outgroup on humanness in a paired sample t-test expecting a small effect size (*d*=.25) and a normal significance level (α=.05), sufficient power (=.80) would be guaranteed with a sample of 128 participants. Therefore, the same number of participants (*N*=130) as in Enock and colleagues (2021) was gathered. Participants were selected using the prescreen function of Prolific Academic to assure that they were all English native speakers and UK residents. They received a mean of £9.84 per hour in Study 1a and £5.98 per hour in Study 1b. Participants who missed more than one attention check question were excluded from the analyses (*N* = 3 in both Study 1a and 1b). The final sample in Study 1a (84 female, 41 male, 2 non-binary) was aged between 18 and 73 (M_{age}=39.23, *SD* = 14.00), while the sample in Study 1b (108 female, 17 male, 2 non-binary) was aged between 19 and 81 (M_{age}=41.47, *SD*=15.24).

2.1.2. Materials

In the current set of studies, the methodology that was proposed by Vaes and Paladino (2010) was adopted, while keeping the experimental materials as similar as possible to that of Enock ad colleagues (2021). Therefore, we used all the 60 traits that were pretested by Enock and colleagues asking participants to judge these traits on their typicality in describing the ingroup ("Please indicate the extent to which you believe each trait word is typical of the average UK national" on an 11-point rating scale ranging from 1="*Not at all*" to 11 "*Very much so*") and the outgroup ("Please indicate the extent to which you believe each trait word is typical of the average immigrant" on an 11-point rating scale ranging from 1="*Not at all*" to 11 "*Very much so*"), their valence ("Please indicate how you

regard every trait from "Very Negative" to "Very Positive" on an 11-point rating scale ranging from 1="*Very Negative*" to 11 "*Very positive*"), and their human uniqueness (in Study 1a; "Please indicate how much the word in each of the following questions applies to humans compared to other species" on an 11-point rating scale ranging from 1="*Only other species*" to 11 "*Only humans*") or their human nature (in Study 1b; "Please indicate how much the word in each of the following questions applies to humans compared to robots" on an 11-point rating scale ranging from 1="*Just Robots*" to 11 "*Just humans*"). The wording of the first two questions and the end points of the ratings scales were directly taken from studies 2a and 2b, while the wording of the latter three questions and their end points were taken from the pretest (see Enock et al., 2021).

2.1.3. Procedure

The questionnaire was conducted online using Qualtrics and participants were contacted through Prolific Academic. Participants were informed that the study aimed to better understand the traits that are typically used to describe UK citizens and immigrants in general. After giving informed consent they were asked some brief demographic questions (age, gender, and nationality) and taken to the trait questionnaires. They responded to the four questions (ingroup typicality, outgroup typicality, humanness and valence) on separate pages for all of the 60 traits. Both the order in which the questions and the traits appeared was randomized between participants.

2.1.4. Design and data analysis

We followed the pre-registered analysis plan calculating within-participant correlations between ingroup typicality and humanity ratings and outgroup typicality judgements and the same humanity ratings in both studies. In this way, a positive correlation indicates that the more a trait is seen as typical in describing a group, the more it

is seen as human, while a negative correlation indicates that mostly traits low in humanness were seen as very typical of that group. We controlled for valence by partialling it out of these correlations; furthermore, in order to normalize distributions, we transformed these correlations to Fisher's Zs (see Michela, 1990). For ease of interpretation, we transformed back and report mean Fisher Z values as correlation coefficients in the tables, while Z-transformed correlation coefficients are depicted in the Figures. We then compared ingroup and outgroup humanness judgments using a paired-sample t-test. In a similar vein and to create an index of intergroup preferences, we also analyzed the within-participant correlations between typicality and valence judgments comparing ingroup and outgroup valence judgments with a paired-sample t-test. It is important to note that within-participant correlations tend to be small as they are based on less stable single judgments compared to group level correlations that are based on group means. In contrast, within-participant correlations provide more power because they are based on a larger number of observations. As a result, small correlations can show to be statistically different from zero.

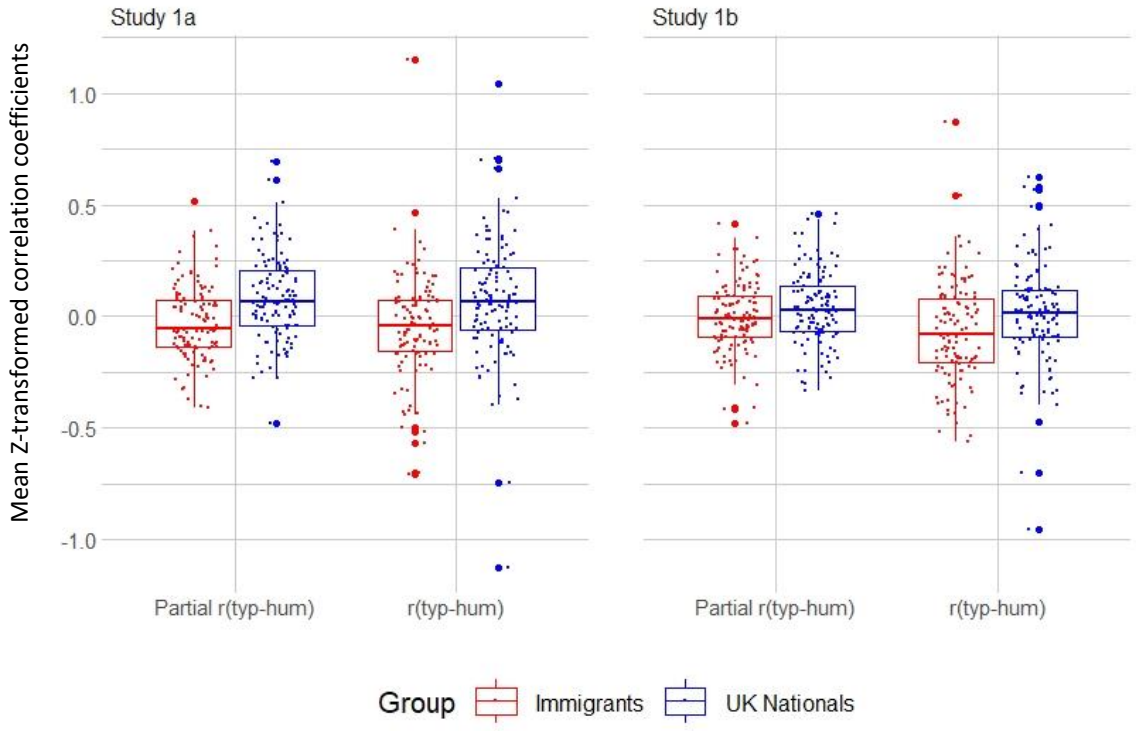
For exploratory purposes, and in order to test the hypothesis that dehumanization effects could be found for both positive and negative traits, we calculated within-participant correlations between typicality and humanness ratings for positive and negative traits separately. The resulting four indices were analyzed in a 2 (Group: Ingroup vs. Outgroup) X 2 (Valence: Desirable vs. Undesirable) within-subject ANOVA.

2.1.5. Results

The humanity that was attributed to UK nationals and immigrants was compared conducting a paired-sample t-test on the within-participant correlations both partialling out valence or not. Due to the fact that some participants always responded the same number on some of the subscales, for 12 participants in Study 1a and 4 participants in Study 1b, no

within-participant correlation could be calculated. Still, we had enough power (.80) to detect a small effect in both studies (*d*=.26). In Study 1a, the paired-sample t-test was significant when controlling for valence, *t*(114)=5.48, *p*<.001, *d*=.68, and when valence was not controlled for, *t*(114)=4.89, *p*<.001, *d*=.60. The same effects were found in Study 1b (*t*(122)=2.22, *p*=.03, *d*=.28, when valence was controlled for and, *t*(122)=3.04, *p*=.003, *d*=.37, when valence was not partialled-out). In both cases, the more a trait was judged to be typical in describing the ingroup the more it was seen as human, while a reverse tendency was observed for the outgroup. As can be seen in Figure 1, this was true for both dimensions of humanness and regardless of whether valence was controlled for or not.

Figure 1: Mean Z-transformed partial and full within-participant correlations between typicality (typ: ingroup vs. outgroup) and humanity judgments (hum: HU in Study 1a, left and HN in Study1b, right). Boxes represent confidence intervals around the median.

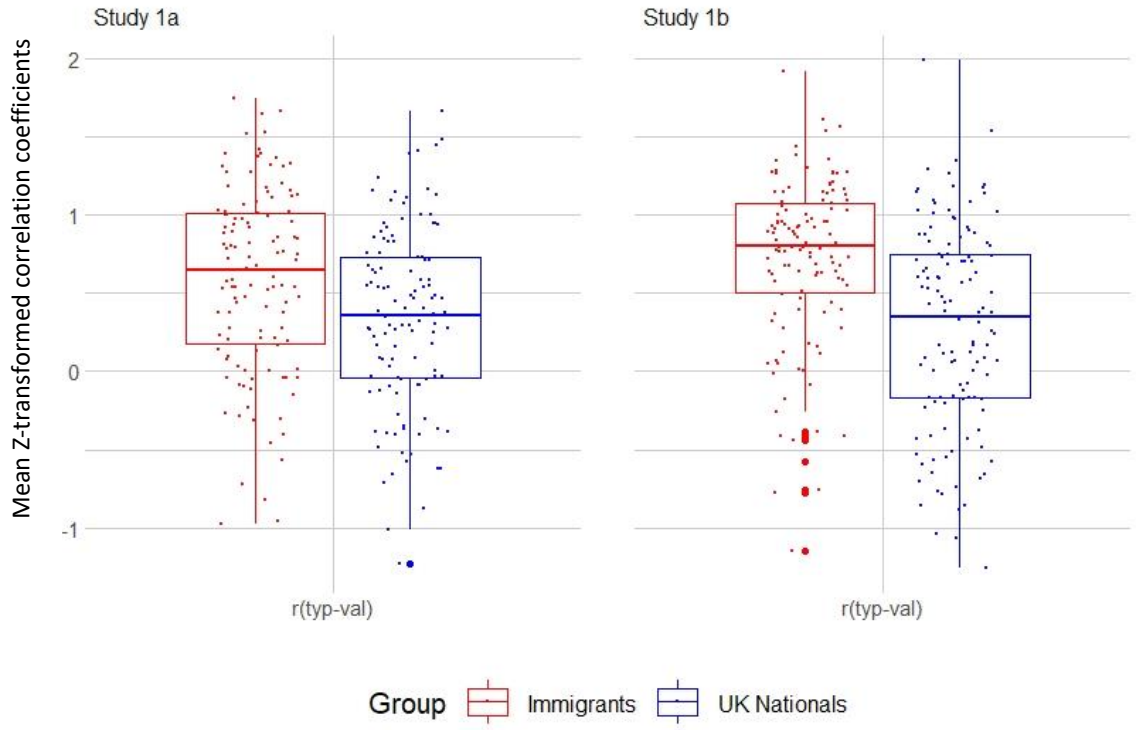


Mean *r*	-.03	.08***	-.05*	.07**		-.003	.04*	-.07***	.01
SD	.16	.19	.23	.27		.15	.16	.22	.23
N	118	118	118	118		123	125	123	125

*Note. All means represent Fisher Z's that are reverted back to rs. Marked means are significantly different from 0, *** p < .001, ** p < .01, * p < .05*

We also verified the valence that was attributed to the in- and the outgroup in both studies calculating a paired sample t-test on the within-participant correlations between typicality and valence judgments. In both studies, this analysis indicated that the outgroup was seen as more positive than the ingroup ($t(114)=-3.33$, $p=.001$, $d=.38$, for Study 1a and, $t(122)=-5.71$, $p<.001$, $d=.55$, for Study 1b; see Figure 2).

Figure 2: Mean Z-transformed within-participant correlations between typicality (typ: ingroup vs. outgroup) and valence (val) judgments in Study 1a (left) and Study1b (right). Boxes represent confidence intervals around the median.



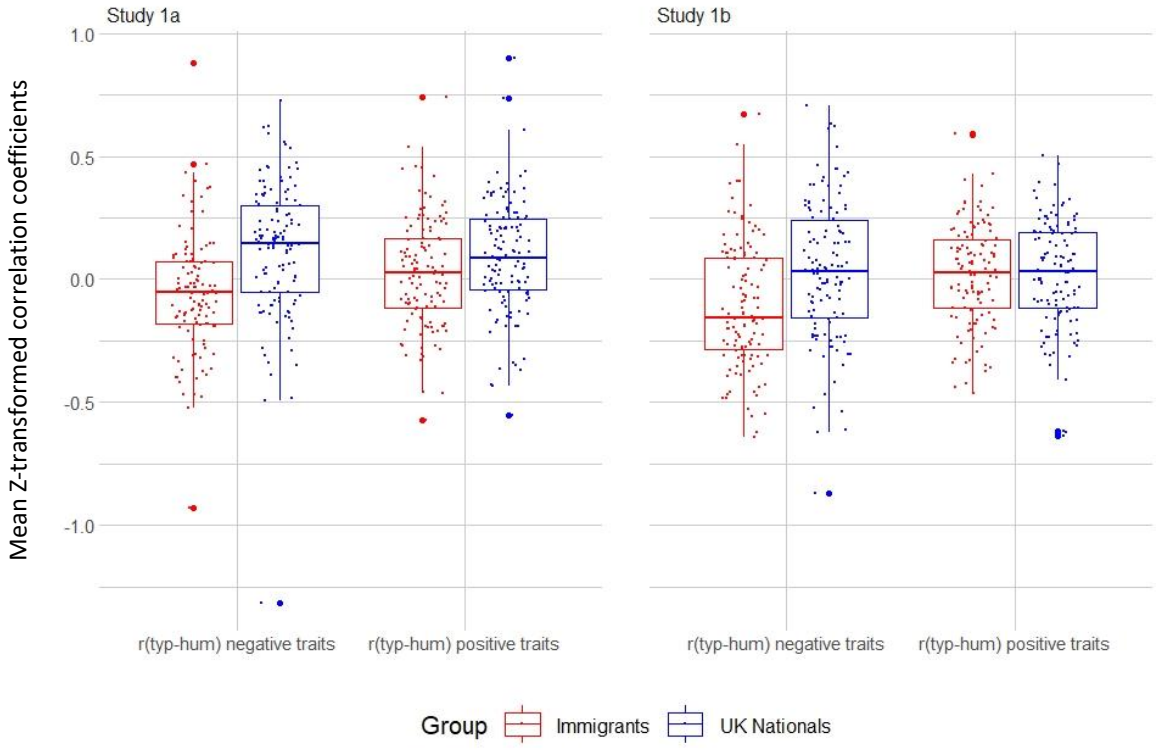
Mean *r*	.53***	.33***		.62***	.28***
SD	.59	.58		.54	.65
N	118	118		123	125

*Note. All means represent Fisher Z's that are reverted back to rs. Marked means are significantly different from 0, *** p < .001, ** p < .01, * p < .05*

Finally, we calculated the within-participant correlations separately for positive ($N=33$) and negative ($N=27$) traits. This analysis allowed us to verify whether the dehumanization effect holds for both positive and negative traits. We analyzed these

within-participant correlations in a 2 (Group: ingroup vs. outgroup) X 2 (Valence: positive vs. negative) within-participants ANOVA. This analysis resulted in the expected group main effect, $F(1, 114)=29.74$, $p<.001$, $\eta^2_p=.21$ in Study 1a and, $F(1, 121)=8.23$, $p=.005$, $\eta^2_p=.06$ in Study 1b, that was qualified in both cases by the two-way interaction [$F(1, 112)=10.92$, $p=.001$, $\eta^2_p=.09$, and $F(1, 121)=15.13$, $p<.001$, $\eta^2_p=.11$, in Study 1a and 1b respectively]. Even though there was an overall tendency to attribute less humanness to the outgroup compared to the ingroup, this difference showed to be significantly stronger for the negative, compared to the positive traits (see Figure 3). Said differently, the more UK nationals described their ingroup with negative traits the more these traits were perceived to be human, while the reverse tendency was true for the outgroup.

Figure 3: Mean Z-transformed within-participant correlations between typicality (typ: ingroup vs. outgroup) and humanity judgments (hum: HU in Study 1a, left and HN in Study1b, right) for positive and negative traits separately. Boxes represent confidence intervals around the median.



Mean *r*	-.06**	.09***	-.0007	.05*		-.11***	.03	.02	.01

SD	.24	.28	.21	.23		.25	.28	.21	.22
N	118	116	118	118		122	122	123	125

*Note. All means represent Fisher Z's that are reverted back to rs. Marked means are significantly different from 0, *** p < .001, ** p < .01, * p < .05*

3. Discussion

Empirical work on intergroup dehumanization over the last 20 years has consistently demonstrated that ingroup members are often perceived to be more human than outgroup members regardless of intergroup prejudice or evalutations. Still, Enock and colleagues (2021) challenged these findings demonstrating that participants tended to attribute desirable human traits more strongly to ingroup members, but undesirable human traits more strongly to outgroup members. As such, these authors claimed that the trait-based dehumanization of the outgroup did not hold when intergroup biases in terms of valence were taken into account.

The current set of studies rebuts these arguments in several ways. Using a more parsimonious methodology to gauge dehumanization including all the traits proposed by Enock and colleagues (2021) and asking participants to judge them on their typicality, humanity and valence, allowed to show an overall tendency of UK nationals to dehumanize immigrants. This effect was observed on both humanness dimensions (HU and HN) when controlling for the valence of the traits or not. These results are in line with previous research that has shown similar findings with a similar control for the valence of the traits in several intergroup contexts including various national (e.g., Albanians, Brazilians, Cubans, Japanese, Americans), regional (e.g., Northern vs. Southern Italians), and ethinic groups (e.g., Gypsies, Slavs; see Vaes et al., 2010; Vaes and Paladino, 2010).

It is interesting to note that the size of the dehumanization effect differed between the two dimensions of humanness. While moderate effects were observed for animalistic

dehumanization, small effects sizes were reported for mechanistic dehumanization. Previous research has mostly reported animalistic forms of dehumanization especially towards low status groups, while mechanistic forms of dehumanization, to my knowledge, have never been reported towards immigrant outgroups.

The dehumanization of the outgroup even emerged when positive and negative traits were analyzed separately. Interestingly, the effect showed to be larger for negative compared to positive traits. When looking at the attribution patterns of participants overall, UK nationals tended to attribute the negative traits to describe their ingroup the more they perceived them to be human, while they attributed negative traits to the outgroup, the more they perceived them low in humanness. While this finding is clearly contrary to the claim of Enock and colleagues (2021), it is in line with previous research of Koval and colleagues (2012) who demonstrated that people tend to humanize the negative characteristics and flaws of their ingroup. Indeed, these flaws seem easier to bear when one can say that they make us “only” human.

Importantly, in the current sample the outgroup was perceived more positively compared to the ingroup. Even though the dehumanization of liked outgroups has been reported before (e.g., Vaes & Paladino, 2010), this was an unexpected finding. It is important to highlight that the sample of Enock and colleagues (2021) only included participants that voted on the political right. Given that the participant pool of Prolific mostly consists of left-wing voters, who generally hold more positive views towards immigrants (e.g., Esses, 2021; Vaes et al., 2015), without imposing any restrictions our sample was probably mostly made up of left-wing voters. To test this possibility, we added a question asking participants’ political affiliation in Study 1b and verified whether it might have moderated any of our findings including the outgroup preference effect. As expected,

only 22 participants indicated to be right-wing voters and only these participants evaluated the ingroup (M_r=.59) more positively compared to the outgroup (M_r=.36), $F(1, 119)=3.99$, p=.048, η^2_p=.03. All left-wing voters, instead, attributed more positive traits to the outgroup compared to the ingroup (see SOM for a full presentation of these analyses). Importantly, political affiliation did not moderate the attribution of humanity to the ingroup and the outgroup. Therefore, sampling differences might explain why participants generally expressed a preference for the outgroup compared to the ingroup in these studies.

All in all, the current manuscript demonstrates that a trait-based account of dehumanization holds value in the study of intergroup bias over and above effects of intergroup evalutations. At the same time and in line with Enock and colleagues (2021), the current work highlights the risks of gauging dehumanization through the attribution of a small number of fixed human traits that are pre-tested in a separate sample and do not allow to control for desirability. Future work on dehumanization should take these recommendations into consideration and conduct their work on a larger set of traits or select their traits carefully, so that valence or desirability effects can be controlled for.

4. Data Reference

Vaes, J. DATA_study1a.sav, Open Science Foundation, 2022, https://osf.io/gwjcv/

Vaes, J. DATA_study1b.sav, Open Science Foundation, 2022, https://osf.io/gwjcv/

5. References

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

Bain, P., Vaes, J., & Leyens, J. P. (2014). *Humanness and dehumanization*. Psychology Press. doi: 10.4324/9780203110539

DeLuca-McLean, D., & Castano, E. (2009). Infra-humanisation of ethnic minorities: The moderating role of ideology. *Basic and Applied Social Psychology, 31*, 102–108. doi: 10.1080/ 01973530902880258

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. https://doi.org/10.1016/j.cognition.2021.104682

Esses, V.M. (2021). Prejudice and discrimination toward immigrants. *Annual Review of Psychology, 72*, 505-531. https://doi.org/10.1146/annurev-psych-080520- 102803

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

Fincher, K. M., & Tetlock, P. E. (2016). Perceptual dehu- manization of faces is activated by norm violations and facilitates norm enforcement. *Journal of Experimental Psychology: General*, *145*, 131–146. https://doi.org/10 .1037/xge0000132

Harris, L. T., & Fiske, S. T. (2006). Dehumanising the lowest of the low: Neuroimaging responses to extreme out-groups. *Psychological Science, 17*, 847–853.

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

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

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.

Koval, P., Laham, S. M., Haslam, N., Bastian, B., & Whelan, J. A. (2012). Our flaws are more human than yours: Ingroup bias in humanising negative characteristics. *Personality and Social Psychology Bulletin, 38*, 283–295. doi: 10.1177/0146167211423777

Kteily, N.S., 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*, 901–931. https://doi.org/10.1037/pspp0000048

Kteily, N.S., & Landry, A.P. (2022). Dehumanization: trends, insights, and challenges. *Trends in Cognitive Sciences, 26(3)*, 222-240. https://doi.org/10.1016/j.tics.2021.12.003

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

Leyens, J-Ph., Paladino, M. P., Rodrıguez-Torres, R., Vaes, J., Demoulin, S., Rodrıguez-Perez, A., et al. (2000). The emotional side of prejudice: The attribution of secondary emotions to ingroups and outgroups. *Personality and Social Psychology Review, 4*, 186–197. doi: 10.1207/S15327957PSPR0402_06

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

Markowitz, D. M., & Slovic, P. (2020). Social, psychological, and demographic characteristics of dehumanization toward immigrants. *Proceedings of the National Academy of Sciences, 117*(17), 9260–9269. https://doi.org/10.1073/pnas.1921790117

Martinez, R., Rodriguez-Bailon, R., & Moya, M. (2012). Are they animals or machines? Measuring dehumanization. *The Spanish Journal of Psychology, 15*, 1110-1122. http://dx.doi.org/10.5209/rev_SJOP.2012.v15.n3.39401

Michela, J. L. (1990). Within-person correlational design and analysis. In C. Hendrick & M. S. Clark (Eds.), *Research methods in personality and social psychology* (pp. 279–311). Thousand Oaks, CA: Sage.

Morris, K.L., Goldenberg, J., & Boyd, P. (2018). Women as animals, women as objects: Evidence for two forms of objectification. *Personality and Social Psychology Bulletin, 44*, 1302-1314. https://doi.org/10.1177/0146167218765739

Paladino, M. P., & Vaes, J. (2009). Ours is human: On the pervasiveness of infrahumanisation in inter-group relations. British Journal of Social Psychology, 48, 237–251. doi: 10.1348/ 014466608X322882

Ruzzante, D., & Vaes, J. (2021). The timeline of mentalization: Distinguishing a two-phase process from mind detection to mind attribution. *Neuropsychologia*, 160:107983. doi:10.1016/j.neuropsychologia.2021.107983

Saminaden, A., Loughnan, S., & Haslam, N. (2010). Afterimages of savages: Implicit associations between 'primitives', animals and children. *British Journal of Social Psychology, 49*, 91-105. doi: 10.1348/014466609X415293

Vaes, J., Heflick, N. A., & Goldenberg, J. L. (2010). ''We are people'': In-group humanisation as an existential defense. *Journal of Personality and Social Psychology, 98*, 750–760. doi: 10.1037/a0017658

Vaes, J., Latrofa, M., Vieno, A., & Pastore, M. (2015). Exposure to politicized media and prejudice against immigrants in Italy: Identifying its impact and psychological mediators. *Psicologia Sociale, 2*, 141-160. doi: 10.1482/80762

Vaes, J., Leyens, J. P., Paladino, M. P., & Miranda, M. P. (2012). We are human, they are not: Driving forces behind outgroup dehumanisation and the humanisation of the ingroup. *European Review of Social Psychology*, *23*, 64–106. https://doi.org/10.1080/10463283.2012 .665250

Vaes, J., & Paladino, M. P. (2010). The uniquely human content of stereotypes. *Group Processes & Intergroup Relations, 13*, 23–39. doi: 10.1177/1368430209347331

Supplemental Online Material

The supplemental online material contains a detailed description of the moderation analyses of participants' political affiliation on the dehumanization and intergroup preference effects that were conducted in Study 1b. These analyses were not pre-registered.

Moderation analyses of participants' political affiliation

In Study 1b, among the demographic questions a single item was added in which participants were asked to move a slider responding to the question: "Some people describe political affiliation on a left to right spectrum. Please indicate where you believe your political ideology lies on this spectrum." The slider was marked with the labels "left", "center" and "right" and could contain a score from 0 (=left) to 100 (=right). We decided to divide the sample into 4 groups based on the quartiles. Given that the 75° percentile equaled 50, only the fourth subgroup was made up of right-wing voters (see Figure S1). We then analyzed the within-participant correlations between typicality and valence judgments in a 2 (Group: Ingroup vs. outgroup) x 4 (Political affiliation: extreme left vs. left vs. center left vs. right-wing voters) mixed ANOVA in which only the latter variable involved a comparison between participants. This analysis resulted in a significant main effect of Group, $F(1, 119)=34.16$, $p<.001$, $\eta_p^2=.22$, that was qualified by a significant interaction with Political affiliation, $F(3, 119)=17.10$, $p<.001$, $\eta_p^2=.30$. This interaction is depicted in Figure S1 and indicates that all left-wing voters judged the outgroup to be more positive than the ingroup, except for the right-wing voters who judged the traits they saw typical in describing the ingroup as more positive compared to those they ascribed to immigrants.

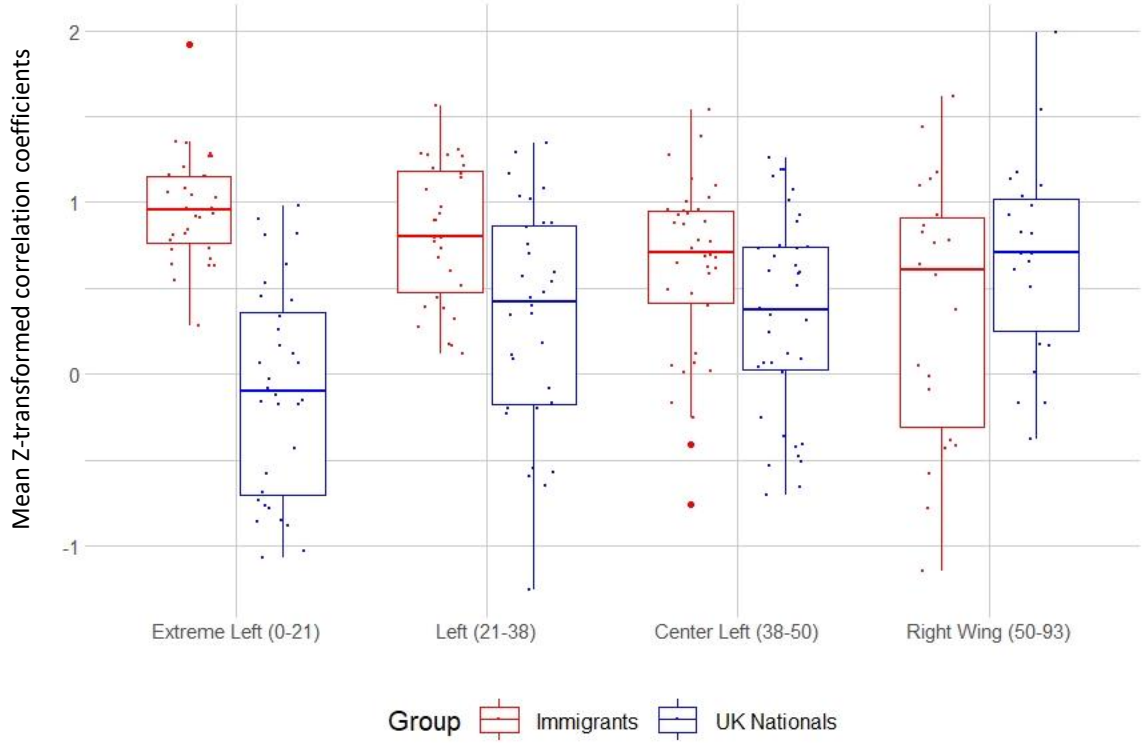


Figure S1: Mean Z-transformed within-participant correlations between typicality (typ: ingroup vs. outgroup) and valence (val) judgments in Study 1b as a function of participants' political affiliation. Boxes represent confidence intervals around the median.

The same analysis was conducted on the partial within-participant correlations between typicality and humanity judgments controlling for valence. From this analysis only a significant Group effect emerged, $F(1, 119)=4.60$, $p=.03$, $\eta_p^2=.04$, indicating that the ingroup was seen as more human than the outgroup regardless of participants' political affiliation.