Commentary on Over (2020): Well-taken Points About Dehumanization, But Exaggerated Challenges

Roger Giner-Sorolla

Pascal Burgmer

Nuray Demir

University of Kent

Final accepted version, *Perspectives on Psychological Science*: June 1, 2020

Corresponding author: Roger Giner-Sorolla, rsg@kent.ac.uk

Abstract

We offer a critical appraisal of Over's (2020) seven challenges for dehumanization research in social psychology. While the challenges mainly attack an exaggerated version of the claims that dehumanization research actually makes, the positive suggestions that follow them open up some much-needed questions. By seeing humanity as a prototype-based category, we can address more sensitively how various departures from prototypical humanity relate to each other. We can also ask whether all of these departures are equally necessary or sufficient explanations for the effects of dehumanizing treatment and metaphors in degrading moral consideration of other groups of people.

Commentary on Over (2020): Well-taken Points About Dehumanization, But Exaggerated Challenges

Over (2020) presents seven challenges to a representation of the prevalent theoretical position about social dehumanization in social psychology. The article concludes with positive suggestions for ways to understand contemporary research and historical examples related to dehumanization. We, as authors, had different initial reactions to these arguments. Still, we ended up agreeing that while the positive suggestions are useful and lead to interesting questions for research, the objections are exaggerated.

Specifically, Over characterizes the target "dehumanization hypothesis" in these words: "1) Victims of intergroup harm are perceived to be similar to non-human entities 2) As a result, natural inhibitions against causing them harm are eroded leading, in extreme cases, to genocide and torture." (page number TBD)

But this is not the statement that the seven challenges actually are aimed at refuting, by and large. Here is our reconstruction of that statement, differences italicized:

1) Victims of intergroup harm, *and nobody else*, are *explicitly* affirmed to be *identical* to non-human entities. 2) *As an inevitable result of dehumanization*, natural inhibitions against causing them harm are eroded, *and all aspects of genocide and torture can be completely explained via this mechanism.*

We know no researchers on intergroup dehumanization who have endorsed such an extreme statement. And yet, Over's objections target the extreme version, but miss many nuances of the more reasonable one. To go through the challenges, one by one:

#1. Examples of ingroups likening themselves to well-regarded animals challenge the extreme claim "victims of intergroup harm, and nobody else …" but not the actual hypothesis that focuses only on victims of harm. Indeed, it is possible for an ingroup to ennoble (or brutalize) itself through comparison to a specific animal, and also to lower and brutalize the enemy group through seeing them as less than human. Nor do the examples refute findings about the *relative* prevalence of ingroup and outgroup dehumanization – most studies of this topic compare the humanity of the ingroup to that of the outgroup, finding a relative difference, without requiring that the ingroup be seen as fully human. And finally, while some individual species may be seen as admirable, the general category "animal" and indeed most animal names are not, we suspect, compliments when applied to people. An exception proves the rule: the generic "animal" becomes grudgingly positive when traits seen as uniquely human, such as inhibition or mercy, must be cast off to achieve some goal ("fight like beasts," "party animals," etc.)

#2 and #4. Showing that victims of intergroup violence are sometimes, even often, described in human (#2) and mentalizing (#4) terms refutes the extreme claim that they are seen as "identical" to animals, but not the actual claim that they are seen as *similar*. In fact, all the research and theory lines reviewed in the paper specify only one or two dimensions of perceived difference as crucial, be it essence, mental development, mental traits, or emotional experience. Here, too, the argument becomes perilously unfalsifiable when explicit descriptions of humans as animals are taken as evidence that the humans are actually *not* seen as animals – "There is no sense in consistently reminding a rat that it is, in fact, a rat" (p. TBD).

#3. If we reject implicit group-animal/human associations as evidence for dehumanization, this refutes only the phantom claim that dehumanization must be *explicit*.

First of all, some evidence for everyday intergroup dehumanization is explicit or "blatant" in nature (Kteily & Bruneau, 2017). But also, as Over acknowledges in the final section, the function of an associative metaphor is to bring to mind information in line with the metaphor, whether or not its proposition is actually endorsed. If a group of people is associated with apes, this facilitates the same kind of judgments of them as mentally and culturally lesser, more violent, etc. as if the similarity were explicitly believed. Research on implicit associations (e.g., Goff, Eberhardt, et al, 2008) supports this point by connecting such associations to more explicit forms of discrimination. The absence of explicit endorsement, whether sincere or motivated by social presentation, does not undermine the reality of the phenomenon.

#5. Again, the specific nature of most actual theories of dehumanization allows for people to be seen both as the human category "criminal" or "avaricious" and as similar to animals in other ways. In fact, there is no other way to interpret the majority of intergroup dehumanization studies which start from the premise of group *labels* that only apply to humans (e.g., "pedophiles," "terrorists" in Giner-Sorolla & Russell, 2019, or the national and racial labels in many other published studies), making it absurd to hold total dehumanization as a standard. Additionally, dehumanization has seldom been claimed to be the motivation behind *every* act of cruelty. In fact, Bandura (1999) outlines numerous other mechanisms of moral disengagement that can both facilitate and justify post-hoc such acts, such as denying responsibility or using euphemistic language. These mechanisms disengage the conscience in many more ways than putting victims outside of moral concern by viewing them as less than human.

#6. To the extent that this objection is distinct from #5, it points out that some forms of cruelty and discrimination are aimed at harming people through their human faculties of imagination, self-regard, etc. This objection, too, aims for the phantom point that total dehumanization is necessary to explain all facets of crimes against humanity. But unlike the previous ones, this challenge does successfully hit an assumption of many theories that is not explicitly tested: the assumption that human-like mentalization is a sufficient, active ingredient to activate the special moral concern that attaches to the human species. It questions, in other words, whether all dimensions of perceived humanity bear equally on moral treatment.

#7. Like #6, this challenge only argues for a plurality of grounds for human status, as indeed the theories reviewed provide, rather than against the importance of human status *per se*. But it does succeed in questioning whether human mental traits are necessary to the moral consideration that attaches to the human category. Going even beyond the provided example of a baby, the moral status of a human fetus or embryo critically depends on whether it is seen as a human life. It is this essential interpretation of human "similarity" that often intervenes in morality, to arouse concern for individuals whose mental status is not at all prototypical.

Although these seven lines of attack mostly miss the point of actual dehumanization theory and research, the positive mechanisms proposed nearer the end of the article are welcome. They demand clarity in a field of research that often confuses several concepts under the label "dehumanization." We now share our thoughts on two more specific implications for research going forward.

1. **Dehumanization as a matter of prototype (de)categorization.** It never was plausible to anyone that human beings are dehumanized by being completely excluded from the human category, or by being treated in every respect like animals. Rather, a more useful concept is that groups of humans might be judged as non-prototypical, metaphorically likened to objects or animals *less* worthy of moral consideration (rather than to angels or lions) and therefore subject to inferior treatment. This corresponds to the dominant *prototype* models of categorization in cognitive psychology (e.g., Rosch, 1975), in which decisions operate via similarity and centrality of exemplars to a set of prototypical traits, rather than via all-or-none categorization.

For example, even though antisocial characteristics such as cunning and villainy may be seen as typical of humans more so than animals, they may also indicate that their possessors are not prototypical humans, and therefore lacking a number of other typical human characteristics that Haslam (2006) and Leyens and colleagues (2000) elaborated in research, such as conscience or capacity for reflection. The prototype model leads to caution about equivalence between these human-defining traits, because a person can be seen as human on one of the following dimensions, yet dehumanized on others:

- Human essential category (vs. all nonhumans; as measured by blatant identification as human, and by associations to the human concept)
- Human mental complexity (vs. animals; as measured by human uniqueness traits and emotions)
- Human (and animal) capacity for spontaneity and creativity (vs. inanimate objects and machines; as measured by human nature traits)
- Human (and animal) capacity for suffering (vs. insensate things)

- Human prosociality (vs. villains or “demons”)

It would be an empirical question whether any given non-prototypical status would entail degradation along other dimensions of humanity. That is, a “criminal,” while acknowledged as having some amount of mind and intentionality, may or may not be seen as similar to non-human entities such as animals (e.g., wild dog), intelligent but emotionless machines (e.g., the Terminator), or more fanciful examples of antisocial nonhumans such as demons, zombies, or monsters. This also calls for empirical tests, rather than assumptions, of whether each departure from prototypical humanity undermines the granting of positive moral rights (e.g. citizenship) and negative ones (e.g. freedom from harm). In this regard, projects such as Bastian, Laham et al. (2011), linking lower levels of different prototypical human traits to differential diminutions of moral responsibility and suffering, are admirable if by no means completely resolved.

But we also should not lose sight of the requirement that to be meaningfully called “dehumanization,” all these dimensions should be demonstrably linked to a reluctance to include target individuals and groups in the human essential category, the first one we have listed. Otherwise, in line with Over’s suggestions, it would be more parsimonious simply to speak of denial of emotional experience, mental capacities, and so on, rather than to use the blanket term “dehumanization.”

If the different elements of the human prototype were shown to be inherently connected to each other, then even applying positive animal metaphors to the ingroup would carry added drawbacks. Perhaps the purpose of these is to dehumanize the ingroup in selected, very specific ways – to promote and advertise the freedom from remorse and inhibition ascribed to animals in the specific context of winning a sports game or a war. Perhaps, all the same, these seemingly complimentary metaphors could create associations

of the ingroup with inhumanity that have paradoxical effects, such as lowering the impulse to cooperate with other ingroup members.

2. Dehumanization can only be partial in a "humanized" world.

Under a prototype model of the human category, the multiplicity of mental dimensions makes less likely the extreme case of dehumanisation in which perceivers deny victims their mind (or "human essence") altogether and only view them as bodies or meat: for example, when consumers perceive animals that they eat to possess a less complex mental life than other domestic animals, such as pets (Bastian, Loughnan, et al., 2012). There are more examples for such denials of minds during wartime (e.g., militaries referring to a "body count" to keep track of the number of killed enemy soldiers). However, even a focus on the physical body does not often lead to full de-mentalization, but rather to a differential perception of certain mental states (Gray, Knobe, et al., 2011). Even in seemingly extreme cases of dehumanisation, people still do perceive minds and mental characteristics in victims.

Indeed, most cases of dehumanisation are probably the result of some kind of flexible mind perception and ascription, as Over points out. After all, humans *are* animals. As Over rightly states, comparisons to non-human entities are not reserved for outgroups (challenge #1, p. TBD). Consequently, dehumanisation seems to rest on the more fundamental phenomenon of ascribing minds to non-human entities in the first place, a phenomenon referred to as anthropomorphism (Epley, Waytz, & Cacioppo, 2007). Children (Hood, Gjersoe, & Bloom, 2012) and adults (Forstmann & Burgmer, 2015) view minds (both of animals and humans) as distinct from the physical world, and we can't help but see them everywhere.

People therefore seem rather inclined to impose at least some dimensions of humanity, rather than neglect it. However, Over writes that “It would be extremely unusual, and most likely inappropriate, to describe an animal or a machine as disloyal or spiteful, for example” (p. TBD). But who has never yelled at their computer or another piece of malfunctioning equipment that was (presumably intentionally) “messing with us”? Similarly, who hasn’t laughed at one of the countless videos on YouTube showing how loyal dogs are, and that cats could not be any more disloyal and indifferent? People are apparently more likely to rely on the idea of lesser minds, or lesser persons, than to completely abandon mind and person concepts when dealing with nonhuman -- or dehumanized -- beings.

To conclude, we think the field of dehumanization research has been overdue for a focus on its fundamental terms and assumptions. Although the challenges presented here by Over do not destroy the evidence for dehumanization, they do force its proponents to be more clear about what they mean, and to do more work to test the coherence of the human category’s many facets. A dimension that forms part of the human prototype may or may not connect directly to categorizing those who lack it as “less human” in essence. And given the ubiquitous occurrence of mind perception, it may be hard to see any sufficiently complex agent as completely lacking in all human mental qualities.

References

Bandura, A. (1999). Moral disengagement in the perpetration of inhumanities. *Personality and Social Psychology Review*, *3*(3), 193-209.

Bastian, B., Laham, S. M., Wilson, S., Haslam, N., & Koval, P. (2011). Blaming, praising, and protecting our humanity: The implications of everyday dehumanization for judgments of moral status. *British Journal of Social Psychology*, *50*(3), 469-483.

Bastian, B., Loughnan, S., Haslam, N., & Radke, H. R. M. (2012). Don't mind meat? the denial of mind to animals used for human consumption. *Personality and Social Psychology Bulletin*, *38*(2), 247–256. https://doi.org/10.1177/0146167211424291

Epley, N., Waytz, A., & Cacioppo, J. T. (2007). On seeing human: A three-factor theory of anthropomorphism. *Psychological Review*, *114*(4), 864–886. https://doi.org/10.1037/0033-295X.114.4.864

Forstmann, M., & Burgmer, P. (2015). Adults are intuitive mind-body dualists. *Journal of Experimental Psychology: General*, *144*(1), 222–235. https://doi.org/10.1037/xge0000045

Giner-Sorolla, R, & Russell, P. S. (2019) Not just disgust: Fear and anger also relate to intergroup dehumanization. *Collabra: Psychology, 5* (1). doi:10.1525/collabra.211

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*, 292-306.

Gray, K., Knobe, J., Sheskin, M., Bloom, P., & Barrett, L. F. (2011). More than a body: Mind perception and the nature of objectification. *Journal of Personality and Social Psychology*, *101*(6), 1207–1220. https://doi.org/10.1037/a0025883

Haidt, J. (2001). The emotional dog and its rational tail: A social intuitionist approach to moral judgment. *Psychological Review*, *108*(4), 814–834. https://doi.org/10.1037/0033-295X

Hood, B., Gjersoe, N. L., & Bloom, P. (2012). Do children think that duplicating the body also duplicates the mind? *Cognition*, *125*(3), 466–474. https://doi.org/10.1016/j.cognition.2012.07.005

Kteily, N.S., & Bruneau, E. (2017). Darker demons of our nature: The need to (re-)focus attention on blatant forms of dehumanization. *Current Directions in Psychological Science, 26,* 487-494. doi:https://doi.org/10.1177/0963721417708230

Rosch, E. (1975). Cognitive representations of semantic categories. *Journal of Experimental Psychology: General*, *104*(3), 192.