

This is my ([Michael Aird's](#)) very unpolished notes/commentary on [Moral expansiveness: Examining variability in the extension of the moral world](#), by Crimston et al.

Here's the paper's abstract:

The nature of our moral judgments-and the extent to which we treat others with care-depend in part on the distinctions we make between entities deemed worthy or unworthy of moral consideration-our moral boundaries. Philosophers, historians, and social scientists have noted that people's moral boundaries have expanded over the last few centuries, but the notion of moral expansiveness has received limited empirical attention in psychology. This research explores variations in the size of individuals' moral boundaries using the psychological construct of moral expansiveness and introduces the Moral Expansiveness Scale (MES), designed to capture this variation. Across 6 studies, we established the reliability, convergent validity, and predictive validity of the MES. Moral expansiveness was related (but not reducible) to existing moral constructs (moral foundations, moral identity, "moral" universalism values), predictors of moral standing (moral patiency and warmth), and other constructs associated with concern for others (empathy, identification with humanity, connectedness to nature, and social responsibility). Importantly, the MES uniquely predicted willingness to engage in prosocial intentions and behaviors at personal cost independently of these established constructs. Specifically, the MES uniquely predicted willingness to prioritize humanitarian and environmental concerns over personal and national self-interest, willingness to sacrifice one's life to save others (ranging from human out-groups to animals and plants), and volunteering behavior. Results demonstrate that moral expansiveness is a distinct and important factor in understanding moral judgments and their consequences.

Overall, I found the paper quite interesting and useful.

Some quibbles

Species vs individuals

The authors gave participants a measure:

assessing financial donation was created for nonhuman animals and the environment. This measure required participants to imagine they had recently inherited a large sum of money. Participants were then asked how likely they would be to donate a portion of it to a range of six charities (e.g., saving chimpanzee habitats, restoring blue-fin tuna populations, saving endangered plant species

The authors found that a participant's "moral expansiveness" predicted their responses on that measure.

They don't note anywhere that each of those three example charities (and perhaps the other three?) "help" *species* more clearly than helping individuals from those species.[1] It seems unclear whether those efforts would be positive, negative, or roughly neutral for individuals from those species (I'm aware there's been a bunch of thought on that by EAs, but I haven't read much of it). It also seems unclear what the effects would be for individuals from *other* species, and thus what the net effects on wellbeing would be.

From memory, I don't think the authors make claims that directly conflict with the points I make just above. But I suspect most readers would think there's a clear connection between expansive moral boundaries, in the sort of sense Peter Singer talks about, and likelihood of donating to those charities. Whereas I would say there isn't. Which, to be honest, actually makes it substantially more interesting to me that there was the correlation the authors' found.

(I think this points to a more general bit of flawed reasoning that people in the EA community have admirably highlighted, rather than being an oversight these authors are *especially* guilty of.)

Relatedly, elsewhere, they use a measure which they describe as measuring, in part, "concern for nonhuman animals (e.g., "a—protecting the habitats of chimpanzees and the other great apes around the world," vs. "b—ensuring the cost of living remains stable in America)". But again, protecting habitats sounds more clearly related to species flourishing than individual flourishing.

Moral reasons not intrinsically about the target population

The latter measure also measures, in part "concern for the environment (e.g., "a—protecting the world's remaining old-growth forests" vs. "b—preventing another U.S. recession")."

I care to a substantial degree about e.g. reducing deforestation. (Though it's not a top priority for me, and if I was planning to act on it in major ways I'd probably first read up more about the wild animal suffering angle.)

But I assign *very* low credence to the idea that plants have intrinsic moral value. My concern is instead about climate change, which in turn is concerning due to its impacts on humans and, via [existential risks](#), potentially anything else we could ultimately realise might be valuable.

From memory, I don't think the authors ever note that participants' responses might have been informed by things like indirect effects, rather than concern about the entities directly focused on in the questions given. Similarly, perhaps some people were motivated to protect chimpanzee

habitats or save endangered tuna species partly or entirely because they like the idea of those things being the case, or humans getting to see them, or whatever, rather than because of concern for those entities themselves.

Again, I don't think this makes the authors wrong about any specific statements. And it may make their findings more interesting/surprising, in a sense. But I'd say it's worth mentioning.

Similarly, another question they ask relates to people's willingness to donate to save charity workers or to save murderers. I see both types of people's experiences as probably at least roughly equally *intrinsically* morally valuable. But I think I'd be notably more willing to donate to save charity workers than murderers, for instrumental reasons.

Self-report

A *lot* of the results are self-reported, but are described as if actual behaviours were observed. Self-report is valuable and cost-effective, but actual observations are usually better, and definitely noting when something is self-report is good. To be fair, this is more like a common problem with papers in psychology (my home discipline, which I do mostly feel fondly towards), rather than with these authors specifically.

To be fair, study 6 does have actual behaviour. But it's a *very* small behaviour, done *within the lab, right after* a survey that's clearly about something like moral boundaries. This is very common in psych and often hard to avoid. But it seems worth the authors explicitly noting as a limitation for generalising from the study, and from memory I don't think the authors explicitly noted that.

Allow qualitative responses when something is weird

They give participants a self-sacrifice scenario. It's somewhat convoluted, and very unusual. I can imagine people interpreting it in a range of different ways, and/or giving an answer for quite a different reason to the reason the author would guess that answer was given.

In such cases, it seems to me (perhaps as received wisdom from my Honours Supervisor) that it's very good practice to provide a box where people can give qualitative explanations of their answers, if they wish, so you can pick up on unexpected interpretations or rationales. (An example from an EA-run survey where this appears to have been done beneficially is [here](#).)

The authors either didn't do this, or didn't mention having done it or what the results were.

Misc thoughts

Lack of explicit discussion of moral theories or empirical beliefs

The authors write:

Overall, these findings provide further evidence that the depth and breadth of people's moral boundaries hold important implications for decision making. Existing theories of moral decision making and action have focused on the role of emotion versus rational deliberation (Batson, 1987; Haidt, 2001; Turiel, 1983), the divergent nature of our moral intuitions (Haidt, 2012), the importance of morality in our self-conception (Aquino & Reed, 2002), and the dyadic nature of morality and mind perception (K. Gray & Wegner, 2011). Moral expansiveness shows that, in addition to these factors, the extent to which we are expansive in granting moral rights is uniquely influential.

It seems to me that two other potentially very important factors are:

- what moral theory (or "folk version", but beyond just "moral intuitions") a person subscribes to
- a person's empirical beliefs about things like probability of animal sentience, how instrumentally useful forests are, whether insect farming is wise for climate change reasons, etc.

I'd hope that theories of moral decision making and action have explicitly addressed those factors. But I guess it's possible none have, in which case I can't fault the authors for not mentioning them in that sentence.

Moral boundaries may be best thought of as multidimensional

The authors write:

In our studies, there was a relatively consistent order of moral priority: family, friends, and in-group members were seen as relatively central, whereas out-group members and nonhuman targets were seen as relatively distal. This need not mean that people move along this continuum of moral concern in a uniform manner, and some individuals may give particularly high concern to some normatively distant entities, such as granting greater moral concern to the environment than to out-group members. Thus, an expansive moral world could take multiple forms, and there may be particular patterns held by different sections of the community (e.g., as a result of cultural differences, dogs may vary from anthropomorphized companions to a food source). We expect that

nuances in moral priority may produce additional insights in terms of the predictive utility of the MES (e.g., human vs. nonhuman decision making).

I think this is an interesting point, and I hope someone follows up with some psychological, empirical research on that. (I'm currently working on a minor "philosophical"/"theoretical" post on that, and similar work was done by [at least one prior post](#).)

I think that, from a theoretical perspective, it *definitely* makes sense to think of moral boundaries as multidimensional (as I'll argue in that post). With that in mind, it's interesting how strong the [internal consistency](#) ("alpha") in this study was. It seems that, for these participants, moral boundaries did effectively operate somewhat as if "one dimensional", in at least one important sense - whether a person's moral boundaries covered a particular type of entity was correlated with whether they covered another type of entity. (Roughly speaking; this is unpolished.) And this seems especially interesting given that some of the "outer" entities were almost certainly non-sentient (the environment), so it's not as if the Peter Singer story should explain (or "justify") all of this correlation.

I'd hypothesise that there'd be less internal consistency on this scale for people who know or have thought more about morality or moral philosophy. E.g., I'd expect that EAs would be much harder to pin down as having "expansive" or "not expansive" moral boundaries on this scale; for example, I'd guess that, compared to the general public, EAs would tend to *intrinsically* value animals more and the environment less (though EAs may also, on average, care unusually much about climate change for *instrumental* reasons).

From memory, I think a similar pattern occurs with political attitudes, in that people who know less about politics (not just current events but also things like how many seats in Congress there are) can be more easily pinned down as having all of their attitudes in one particular part of the political spectrum than people who know more about politics. (But this is something I remember learning in 2017, and haven't checked since then.)

I think it'd be cool for someone to test that hypothesis in the case of moral boundaries. (Though I say that more out of curiosity than based on EA principles.)