A more thought-ful ape*

Comments on *A Better Ape* by Victor Kumar and Richmond Campbell

Mara Bollard

Weinberg Institute for Cognitive Science
University of Michigan

Introduction

In *A Better Ape,* Victor Kumar and Richmond Campbell (2022) provide an ambitious and compelling history of the evolution of human morality. Informed by evidence from an impressively vast multidisciplinary literature, they offer a rich bio-cultural evolutionary explanation of how the human moral mind arose and developed over time that has wide appeal for philosophers and scientists alike. The descriptive project, which unfolds in the first three parts of the book, would be a tremendous contribution all on its own, but Kumar and Campbell don't stop there: in the final part of *A Better Ape*, Kumar and Campbell shift focus from how human morality has evolved to the squarely ethical question of how humans might become morally evolved. In other words, what would it mean for *Homo sapiens* to become morally *better* apes?

In these comments, I examine Kumar and Campbell's novel moral psychology and raise questions about the relationship between moral norms and core moral emotions. I also argue that social essentialist cognition has a place alongside the moral emotions as a key ingredient in Kumar and Campbell's view of the moral mind as well as in their normative account of how to work towards moral progress.

Moral emotions, norms, and affective resonance

Much of *A Better Ape* is devoted to developing an original pluralist moral psychology. Kumar and Campbell argue that the human moral mind has three ingredients: a set of core moral emotions, a set of core moral norms, and a core capacity for moral reasoning. Kumar and Campbell's story about the origin of human morality begins with a theory of ape morality. Ape morality consists in the *binding* moral emotions: the capacities for sympathy and loyalty (Chapter 1). Humans inherited homologues of sympathy and loyalty, but as human social groups grew in size and complexity, they

evolved new *collaborative* moral emotions of trust and respect, thereby enabling more complex kinds of cooperation. Humans also evolved second-order *reactive* moral emotions, which are elicited when someone fails to exhibit any of the four "basic" moral emotions of sympathy, loyalty, trust, or respect. Kumar and Campbell identify resentment and guilt as the most prominent reactive emotions while also allowing that other emotions such as shame or regret (which, like guilt, are felt toward oneself) also belong in this category (Chapter 2). Together, the binding emotions, collaborative emotions, and reactive emotions comprise the emotional core, which is distinctive to human morality.

Alongside the emotional core, humans evolved a normative core (Chapters 3-4). Kumar and Campbell argue that the first four moral emotions co-evolved with a set of five core moral norms. In brief, sympathy selected for harm norms, loyalty selected for kinship norms, trust selected for reciprocity norms, and respect selected for autonomy norms. Trust and respect joined forces to select for fairness, the fifth moral norm. Kumar and Campbell note that the core moral norms are best understood as open-ended norm *clusters,* or types, rather than "singletons" (Kumar & Campbell 2022, 93). Thus, all the specific norm instances within a norm cluster are unified by their association (or "resonance") with the relevant moral emotion(s).

Though the moral emotions are innate and universal, Kumar and Campbell emphasize that they are also, by evolutionary design, *flexible*. Importantly, the flexibility, or plasticity, of moral emotions explains widespread variation in the elicitation, intensity, and expression of moral emotions. Kumar and Campbell write,

> Whereas everyone in a group has some general dispositions to experience moral feelings like sympathy or trust, a particular situation may trigger sympathy or trust in some people but not in others. Even when everyone does experience moral emotions, they may differ in intensity. Or they may lead to subtly different patterns of behavior (74).

In the above passage, Kumar and Campbell point to differences in the moral emotions of individuals in the same community, but their account also predicts and discusses variation between groups of people due to differences in experience, particularly once social institutions formed and paved the way for diverse institutional moralities (Chapters 6-7). People from different societies across both time and space possess "disparate moral minds:" that is, minds containing moral emotions, norms, and reasons with vastly different *contents,* despite sharing the same emotional and normative core (171).

So, the contents of moral emotions can also differ between societies and across time.

The precise content of moral norms is likewise variable across groups. Consider, for example, fairness norms, which differ widely across cultures. Some cultures privilege egalitarian fairness norms, which will give rise to yet more specific norms and rules about distributing resources equally. Other cultures may instead adopt retributive fairness norms, which dictate that people should be rewarded and punished on the basis of effort, which in turn might lead to unequal resource distribution (93-94).

Kumar and Campbell compare their account with another broadly similar pluralist moral psychology developed by Jonathan Haidt, Jesse Graham, and their colleagues in multiple works (Graham, Haidt, et al. 2013; Haidt 2012; Haidt & Joseph 2004). Any plausible pluralist view of moral psychology should be able to provide a descriptive explanation of why different sociocultural groups have different moral norms and practices, while also adequately explaining what unites diverse moralities under the universal umbrella of human morality, so it's worth briefly examining how the two accounts approach these questions.

Haidt, Graham, and their colleagues seem, at least at first glance, to have ready answers to both of these questions. As Kumar and Campbell put it, "Haidt and Graham think that the *content* of core moral norms is innate" and thus universal (Kumar & Campbell 2022, 94). Haidt, Graham, and their colleagues also think that the implementation or expression of moral norms (also referred to as "moral foundations" in Graham et al. 2013 and Haidt 2012; cf. Haidt & Joseph 2004's discussion of "moral modules") can be flexible and may differ between groups, which in turn allows for significant cross-cultural variation when it comes to more specific norm content. Graham et al. suggest that we should understand the moral mind as a universal and innate "first draft" that gets "filled in and revised… in variable ways across cultures" (Graham et al. 2013, 64-65).

By contrast, Kumar and Campbell deny that the core moral norms themselves are innate. Instead, Kumar and Campbell argue that we have an innate capacity to learn (and be motivated by) norms, plus an innate set of resonant moral emotions. These two capacities are, they claim, all we need to explain the universality — as well as the diversity — of specific moral norms. So, Kumar and Campbell argue that Haidt, Graham and their colleagues are mistaken: there's no need to posit moral norms that are "innately programmed into our brains" (Kumar & Campbell 2022, 94) when the innate moral emotions plus culture can get the job done. This difference explains, at least in part, why Kumar and Campbell take their theory to be preferable to that of Haidt, Graham, and their colleagues.

While we could quibble with aspects of Kumar and Campbell's characterization of the

rival view — for instance, Graham et al. (2013) describe the innateness of the human moral mind as a preparedness "to *learn* values, norms, and behaviors" (63; emphasis added); they also explicitly deny that the moral foundations are discretely localized in the brain (96) — I won't pursue those details here. Instead, I take the apparent dispute between these two accounts to raise some further questions, and a possible tension, for Kumar and Campbell. In particular: given that they wish to avoid positing innate norm contents (and any corresponding biological brain basis), what should Kumar and Campbell say about the *contents* of the core moral emotions, which *are* innate, as well as subserved by dedicated neurological mechanisms? (Kumar & Campbell 2022, 48)

Let's revisit what Kumar and Campbell say about the co-evolution of moral emotions and moral norms:

> Over hundreds of thousands of years, the emotions and norms that lasted were those that helped to resolve problems of interdependent living and that *resonated with each other*. Emotions were inherited vertically by offspring, while norms were inherited vertically and horizontally by students (89, emphasis added).

What does it mean for emotions and norms to "resonate" with each other? Though Kumar and Campbell don't specify precisely what they take the resonance relation to consist in, they say that affectively resonant norms "mesh with" our emotions (87). Additionally, they explain that norms that resonate with, or are concordant with, our emotions are more likely to stick within individual minds and take hold in groups compared to norms that lack affective resonance, and are thus more likely to be passed down via cultural evolution. For example, norms that prescribe helping and forbid harming fellow group members are more affectively resonant than norms that command violence against group members because the former but not the latter resonate with the core moral emotion of sympathy (87-88).

So, norms are more likely to resonate affectively if they command (or forbid) behaviors that typically elicit core moral emotions. Perhaps, then, we could think of specific instances of norm following and norm violation as *characteristic elicitors* of moral emotions. Most philosophers and psychologists (though not all: see, e.g. Shargel & Prinz 2018) who study emotions think that emotions are not only elicited by objects, they are also *about* those objects, in the sense that they *represent* those objects as having some evaluative property: for example, fear represents, or appraises, objects as dangerous, sadness represents objects as involving a loss, and anger represents elicitors as offenses or slights (Lazarus 1991).

The foregoing brings us to a larger question: which theory of emotion do, or should, Kumar and Campbell endorse? While I will not take up the question of whether all of

Kumar and Campbell's candidate moral emotions are in fact emotions,[1] I suspect that a fuller account of the resonance relationship between moral emotions and norms will ultimately require Kumar and Campbell to take a stand on which theory of the emotions, including a characterization of emotions' representational content, is correct.[2]

Why think that Kumar and Campbell need to dive into the question about the representational content of emotions? Recall that Kumar and Campbell hold that many different specific norm singletons, which can vary widely in content between groups and over time, are nevertheless unified as part of the same core norm cluster because they all resonate with the same corresponding core moral emotion(s). I want to suggest that the most natural way of cashing out the resonance relation between norms and emotions — and, relatedly, of explaining how sets of diverse norm singletons converge under particular norm clusters — is to say that it consists in a *content match* ("concordance") between the emotions' representation (or appraisal) of elicitors (i.e., specific instances of norm violations and norm following) and the unifying theme of the relevant norm cluster. In other words, each moral emotion responds to *and represents* specific behaviors as, e.g., fair or unfair, harmful or helpful, autonomy-respecting or autonomy violating, and so on.

Invoking emotions' appraisal contents has another explanatory benefit. As we saw earlier, Kumar and Campbell (rightly) observe that the very same situation can elicit a variety of emotional responses in people within the same group; e.g. some behavior or situation might elicit sympathy or trust in some group members but not in others (Kumar & Campbell 2022, 74). Such interpersonal variation is easy to explain with reference to differences in appraisal contents: the very same elicitor triggers different emotional responses in different people precisely because they have appraised the object differently. Consider the following example: As an Australian, I respond to Vegemite-covered toast with delight, but my American friend finds it disgusting. Why is this? Our respective appraisals of Vegemite contain different contents and thus explain our drastically different emotional responses. It's plausible that our appraisals are influenced by individual factors (e.g., salt tolerance) as well as cultural factors (e.g., what kinds of foods and flavors are common in Australia vs. America; whether yeast extract is considered acceptable as an edible product, etc.).

To get a feel for how the same sort of explanation might apply to cross-cultural variation in moral norms, let's explore fairness norms specifically in the context of the ultimatum

---

[1] For relevant discussion on this point see Odenbaugh, this volume.
[2] It should be noted that in what follows I speak somewhat loosely in that I do not sharply distinguish between two distinct views about emotions: according to the first, emotions are (at least in part) *constituted by* appraisals, whereas the second holds that emotions are *caused by* appraisals. For detailed discussion and comparison of these views, see Scarantino (2010).

game (which Kumar and Campbell discuss in Chapter 3). In the ultimatum game, one player (the giver) is given a sum of money and told they must offer some portion of it to a second player. The second player (the responder), may either accept the offer or reject it: if they reject it, neither player receives any money. Kumar and Campbell note that people in most societies tend to behave in a similar way when playing the ultimatum game: givers typically offer close to half of the money, and responders tend to reject offers that are significantly lower than half. Why is this? According to Kumar and Campbell, "players share a [fairness] norm of equal distribution" (75). Givers are motivated to uphold the norm in part for the sake of equality itself, but also to avoid possible punishment, i.e., having their offer rejected by the responder and subsequently receiving nothing.

Kumar and Campbell's examination of the ultimatum game is understandably brief and primarily serves to inform their subsequent account of why norms and costly punishment evolved. Nevertheless, given Kumar and Campbell's acknowledgement and discussion of widespread diversity when it comes to fairness norms in particular, it is worth exploring human behavior in the ultimatum game a little further.

Joseph Henrich and collaborators (2005; 2006) conducted cross-cultural studies in countries from five continents to examine behavior in economic games, including the ultimatum game, in a diverse range of human social groups. Though the roughly equal strategy mentioned above was favored by some groups, overall the researchers observed significantly more behavioral variation across groups in the ultimatum game than had been previously found when it came to mean offer amounts, rejection frequency, and mean *accepted* offers. To give just one example from their many fascinating results, Henrich et al. (2005) found that Machiguenga responders, members of a small-scale society in Peru, accepted all but one offer despite the fact that more than 75% of offers were lower than 30% of the total money available to the giver.

Henrich et al. (2005) argue that much of the observed behavioral variation in economic games can be explained by group-level differences in economic organization, such as the extent to which economic life depends on cooperating with people outside of one's immediate family and how common market exchange is in a given society. They go on to claim that these economic and social differences may "create between-group differences in *notions of fairness* and punishment [norms]" (807: emphasis added).

This is, to my mind, all perfectly consistent with Kumar and Campbell's account of cultural variation in fairness norms, so Kumar and Campbell can easily agree with Henrich et al. that Machiguenga responders have a notion of fairness that differs from other groups; perhaps one that is less egalitarian and more retributive (Kumar &

Campbell 2022, 93-94). Plausibly, Machiguenga responders simply aren't perceiving low offers as unfair most of the time, and thus are happy to accept these offers.

Further, I suggest that Kumar and Campbell's story could (and should) go even further by explaining cross-culturally variable notions of fairness in terms of differences in how the relevant moral emotions — namely, trust, respect, and resentment — represent and respond to specific behaviors that instantiate or violate fairness norms. They could say that accepted offers in the ultimatum game are perceived (i.e., *appraised*) as fair, and thus elicit feelings of trust and/or respect: i.e., the moral emotions that resonate with fairness. By contrast, unfair offers are appraised as such by the reactive emotion of resentment, which will in turn prompt responders to reject the offer as an expression of punishment.

A final note before I move onto the next section: One might worry that all this talk about emotions' representational content risks over-intellectualizing the emotions in a way that might not be especially friendly to Kumar and Campbell's evolutionary story. In particular, might it prevent them from ascribing innate moral emotions to other apes or to very young children? Fortunately, this need not present a problem. According to many contemporary emotion theorists,[3] contentful appraisals or representations don't have to be understood as articulable propositions that presuppose language mastery or the possession of specific concepts (Scarantino 2010). So, it is still possible for Kumar and Campbell to maintain that the moral emotions are homologous while also developing a more detailed view about the contents of moral emotions, which I've argued they have reason to do.

Language, cognition, and essentialism

Language is widely considered to be a hallmark of our species, perhaps even "our most powerful cultural and cognitive tool" (Perszyk & Waxman 2018, 232). In Part II of *A Better Ape* (Chapters 3-5), Kumar and Campbell discuss the importance of the evolution of language. They argue that the emergence of language not only allowed humans to effectively encode, spread, and preserve useful information, it also played an essential role in the evolution of social reasoning (Kumar & Campbell 2022, 105). In this section, I examine some of Kumar and Campbell's claims about the relationship between language and reasoning, which I take to be broadly compatible with — and further bolstered by — a large body of research in developmental cognitive science which supports the existence of a deep link between human language and cognition.

The link between language and cognition emerges very early in life, and makes possible

---

[3] Perhaps even *most* of them, though not those who are judgmentalists about emotion.

a number of conceptual and representational capacities that distinguish human cognition from the cognitive capacities of other animals (for review, see Perszyk & Waxman 2018 and Waxman 2013). Much of the research exploring the link between language and cognitive development comes from work on early word learning, with a special focus on how naming facilitates object categorization. The evidence supports the existence of a strong connection between naming and object categorization, even before infants have begun to form and use words themselves. For example, when a series of different toy objects (e.g., dog, horse) is introduced to infants one by one along with a novel naming word (e.g. each toy is labeled a "blick" when presented), infants' ability to detect the category-based commonality among the individual objects (e.g. that they are all animals) was uniquely enhanced by the use of naming words (Waxman & Markow 1995; Fulkerson & Waxman 2007).

More broadly, language appears to be crucial to explaining why humans' conceptual and representational capacities are so much greater than those of other species, despite the fact that non-human animals and humans alike are capable of object, numerical, and spatial representations (Perszyk & Waxman 2018). Language enables humans to not just form representations within each of these distinct, encapsulated systems of knowledge, it also enables us to combine representations from otherwise distinct knowledge systems, which in turn allows for higher-order, abstract representations (Spelke & Kinzler 2007; Carey 2009; Spelke 2017). These insights dovetail with Kumar and Campbell's observations that reasoning "requires complex representational systems that other animals lack to the same degree" (Kumar & Campbell 2022, 105). They also set the stage for Kumar and Campbell's account of the uniquely human capacity for social reasoning.

In Chapter 5 of *A Better Ape*, Kumar and Campbell argue that reasoning is thoroughly social. Humans' ability to reason together explains, at least in part, how our human ancestors were able to accumulate so much knowledge about their environment, which in turn helped them to survive. Kumar and Campbell point out that interactive reasoning is, in many cases, more successful than reasoning alone: for example, groups of people working together on the Wason selection task are four times more likely to give the correct answer compared to individuals who tackle the problem alone. So, socially interactive reasoning, particularly within cognitively diverse groups, counteracts individual biases that can lead us to the wrong conclusion if left unchallenged (110-12).

An intriguing and novel part of Kumar and Campbell's account of social reasoning is that its evolution *depended on morality.* They write,

> To be most effective in ferreting out the truth, humans exchanging reasons had

to be motivated by a shared intention to reason cooperatively. This involved seeing others as worthy of respect, as equal partners in a task they value collectively. In this way, natural and cultural selection for interactive reasoning depended on background moral values of equality and moral norms of fairness. Thus, human knowledge rests on interactive reasoning, and interactive reasoning itself rests on morality. Reasoning could not reliably work to produce knowledge unless partners in reasoning were involved in a moral relationship (114).

So, evolution designed reasoning to generate knowledge via social interaction, and interactive reasoning is underpinned by moral emotions and norms that enable humans to see their fellow humans as equals worth reasoning with. However, Kumar and Campbell caution that humans are not, by nature, disposed to view *all* other humans as equal reasoning partners because the core moral emotions are limited in scope. For example, limited sympathy between groups led to outgroup antagonism in ancestral humans. Limited sympathy also, importantly, helps explain modern moral exclusivity more generally. Moreover, limited respect between males and females within groups led (and still leads) to sexism and female subordination (51-52).

In Part IV of the book, Kumar and Campbell extend this explanation to modern examples of humans reasoning and behaving badly. They provide an illuminating account of how the lack of trust and respect between present-day political tribal groups — which is largely driven by socioeconomic inequality — contributes to the pervasive problem of misinformation, particularly when it comes to climate change. Kumar and Campbell convincingly argue that rational discourse and bipartisan pursuit of truth about major moral and political issues requires moral scaffolding — which, unfortunately, is often in short supply in current political and moral discourse (248-250).

To recap briefly: Kumar and Campbell argue that our tendencies toward exclusivity and inequality are driven primarily by the circumscribed nature of our moral emotions. This is all the more troubling once we grasp how failing to treat all fellow humans as equal reasoning partners undermines our ability to reason effectively together. While I find this explanation compelling, I think there is more to the story about how the ingredients of the moral mind lead us morally astray. Here, then, is a supplementary proposal: in addition to stemming from moral emotions that are limited in scope, I suggest that our all-too-human tendencies to exclude and subordinate other people are also shaped to some degree by the evolution and development of essentialist beliefs about social categories.

Psychological essentialism is the view that humans represent certain categories as having an underlying essence. As Eleonore Neufeld succinctly puts it, "we treat

essentialist categories as natural rather than constructed, discovered rather than invented, homogeneous, and as having sharp category boundaries" (Neufeld 2022: 2). Essentialism involves thinking of category members as fundamentally similar to each other (in virtue of sharing the same category essence) and different from members of other categories (who have a different essence). Essentialism also often leads to a tendency to assume that the properties of an individual, even properties that are not directly observable, generalize to other members of that category (Rhodes & Moty, 2020).

Social essentialism refers more specifically to the tendency to think of *social* categories, such as gender, race, and ethnicity, in this way. People who essentialize social categories usually view all members of a particular category as having the category essence, which is thought of as deep, stable (perhaps even immutable), biologically determined, and causally powerful (Gelman 2003).

While researchers disagree about whether essentialist thinking is produced by an innate cognitive module (Atran 1998) or instead develops during childhood as the result of other basic, domain-general processes that support conceptual development (Gelman 2003; Rhodes & Moty 2020), it is nevertheless clear that essentialist thinking emerges as early as three years of age (Gelman 2003; Gelman 2004) and is plausibly considered a universal feature of human cognition that evolved to help us more easily make sense of the world — and of the people in it.

The development of essentialism is closely linked to our capacity for language. For instance, children's essentialist beliefs arise in part from how adults use generic language to describe social categories (Rhodes, Leslie, Bianchi & Chalik 2018). Additionally, just as naming is essential to the establishment of object categories, it is also instrumental in the formation of gender and racial categories, as well as the subsequent use of those categories in inductive reasoning (Waxman 2010).[4]

Though some social essentialist beliefs emerge early on, there is also significant developmental, environmental, and cultural variability in social essentialism. Children develop strong essentialist beliefs about gender by ages 3-5, whereas essentialist beliefs about race tend to develop later and more slowly, and vary much more widely across people and groups (Rhodes & Gelman 2009; Rhodes & Mandalaywala 2017; Pauker, Tai, & Ansari 2020).

---

[4] For example, Sandra Waxman (2010) found that providing 3- and 4-year old children with a novel category name for a target individual (e.g. this one, a white woman, is a "Wayshan") emphasized that individual's membership within a particular social category and subsequently opened up the use of that *kind of person* (i.e. a white person or a woman) as an inductive base for reasoning about other people.

Essentialist representations do not necessarily accurately reflect the structure of the world — race, for instance, is commonly thought of as a social construct rather than a natural or biological kind. Yet, essentialist beliefs are nevertheless widespread across diverse cultural contexts and central to how people perceive and reason about others. The consequences of essentialist cognition are mixed: on the one hand, cognitively sorting people into categories helps us simplify and make sense of a complex social world (Rhodes, Leslie, Bianchi & Chalik 2018). On the other more sinister hand, essentialist thinking about social categories contributes to a whole host of negative attitudes and behaviors, such as increased endorsement of social hierarchies, stereotyping, racial prejudice, discrimination, and decreased desire for cross-group interaction (e.g. Bastian & Haslam 2006; Leslie 2017; Mandalaywala, Amodio & Rhodes 2018; for review, see Pauker et al. 2020 and Neufeld 2022).

With all of this in mind, I contend that social essentialist cognition has an important role to play alongside the moral emotions in Kumar and Campbell's account of the moral mind — and in their explanation of moral exclusivity and inequality. This proposal fits neatly with multiple aspects of Kumar and Campbell's view. For one thing, social essentialism plausibly helps explain why hierarchies often track social group membership. For another, both moral emotions and social essentialist thought display significant contextual and cultural variability.

Relatedly, both moral emotions and essentialist beliefs are, at least to some extent, flexible, and can be shaped by experience — sometimes in ways that contribute to moral progress. Just as one's moral emotions of sympathy and respect can expand in scope the more one spends time with people who were previously outside one's circle of moral concern (Kumar & Campbell 2022, Chapter 9), potentially harmful essentialist beliefs (e.g., about race) decrease the more that people interact with members of social groups that are different from one's own (Pauker et al. 2020).

Further, it appears that group differences in essentialism are a key factor underlying political disagreements. Recent research suggests that a tendency toward essentialist thinking in general is more strongly linked to conservatism than to liberalism. Psychologists Jeremy Clifton and Nicholas Kerry found that conservatives, much more than liberals, endorse the belief that the world is inherently hierarchical, in the sense that it is full of differences that reflect objectively real and important differences (Clifton & Kerry, 2022). In other words, conservatives typically believe that perceived lines between categories — including but not limited to human social categories — matter, and therefore should be preserved. On the other hand, liberals are more likely to see distinctions between categories as superficial or culturally-based, and are therefore less likely to care about maintaining those distinctions.

Clifton and Kerry (2022) describe the difference in hierarchical thinking as *the* main difference between liberals and conservatives. If this is right (and I'm right to cast hierarchical belief in essentialist terms), then we've not only found a way to further elucidate the mechanisms underpinning profound moral and political tribal disagreement, we may also be better equipped to develop specific strategies for bridging such divides. Kumar and Campbell tell us that the key to fostering moral progress (and resisting moral regress) is institutionally scaffolded moral reasoning among diverse communities. I claim that in some cases, this might work, at least to some extent, by decreasing the strength of people's essentialist beliefs. Some suggestive evidence along these lines comes from Rhodes and Gelman (2009), who found that children growing up in racially homogeneous and socially conservative environments tend to think of race as an objective category that sharply distinguishes between different kinds of people. These children also developed increasingly stronger essentialist beliefs about race with age. However, importantly, the same was *not* true of children growing up in more racially diverse and liberal environments.

Of course, there's much more to be said to properly untangle the intricate relationship between essentialist cognition and morality. Still, I think we have reason to be hopeful that a deeper understanding of essentialism can not only fruitfully supplement Kumar and Campbell's already rich view of the moral mind, but also add some more strategies to their playbook for how we might become morally better apes in the future.

Acknowledgements

References

Atran, S. (1998). Folk biology and the anthropology of science: Cognitive universals and cultural particulars. *Behavioral and Brain Sciences, 21*(4), 547-569.

Bastian, B., & Haslam, N. (2006). Psychological essentialism and stereotype endorsement. *Journal of Experimental Social Psychology*, 42(2), 228–235.

Carey S. (2009). *The Origin of Concepts.* Oxford, UK: Oxford University Press.

Clifton, J. D. W., & Kerry, N. (2022). Belief in a Dangerous World Does Not Explain Substantial Variance in Political Attitudes, But Other World Beliefs Do. *Social Psychological and Personality Science*. doi:10.1177/19485506221119324

Fulkerson, A.L. & Waxman, S.R. (2007). Words (but not tones) facilitate object categorization: evidence from 6-and 12-month-olds. *Cognition 105*(1), 218–28.

Gelman, S. A. (2003). *The essential child: Origins of essentialism in everyday thought.* Oxford University Press.

Gelman, S. A. (2004). Psychological essentialism in children. *Trends in Cognitive Sciences, 8*(9), 404–409.

Graham, J., Haidt, J., Koleva, S., Motyl, M., Iyer, R., Wojcik, S. P., and Ditto, P. H. (2013). Moral foundations theory: The pragmatic validity of moral pluralism. In P. Devine and A. Plant (Eds.), Advances in experimental social psychology (Vol. 47, pp. 55–130). Academic Press.

Haidt, J. (2012). *The Righteous Mind. Why Good People are Divided by Politics and Religion.* New York: Pantheon.

Haidt, J., & Joseph, C. (2004). Intuitive Ethics: How Innately Prepared Intuitions Generate Culturally Variable Virtues. *Daedalus*, *133*(4), 55–66.

Henrich, J., McElreath, R., Barr, A., Ensminger, J., Barrett, C., Bolyanatz, A., Camilo Cardenas, J., Gurven, M., Gwako, E., Henrich, N., Lesorogol, C., Marlowe, F., Tracer, D., & Ziker, J. (2006). Costly punishment across human societies. *Science, 312*(5781), 1767-1770.

Henrich, J., Boyd, R., Bowles, S., Camerer, C., Fehr, E., Gintis, H., McElreath, R., Alvard, M., Barr, A., Ensminger, J., Henrich, N. S., Hill, K., Gil-White, F., Gurven, M., Marlowe, F. W., Patton, J. Q., & Tracer, D. (2005). "Economic man" in cross-cultural perspective: Behavioral experiments in 15 small-scale societies. *Behavioral and Brain Sciences, 28*(6), 795–855.

Kumar, V., & Campbell, R. (2022). *A Better Ape: The Evolution of the Moral Mind and How It Made Us Human.* Oxford University Press.

Lazarus, Richard S. (1991). *Emotion and Adaptation*, New York: Oxford University Press.

Leslie, S. J. (2017). The original sin of cognition: Fear, prejudice, and generalization. *The Journal of Philosophy*, 114(8), 393–421.

Mandalaywala, T. M., Amodio, D. M, & Rhodes, M. (2018). Essentialism promotes racial prejudice by increasing endorsement of social hierarchies. *Social Psychological and Personality Science, 9* (4), 461-469.

Neufeld, E. (2022). Psychological essentialism and the structure of concepts. *Philosophy Compass*, 17(5), e12823. https://doi.org/10.1111/phc3.12823.

Pauker, K., Tai, C., & Ansari, S. (2020). Contextualizing the development of social essentialism. *Advances in Child Development and Behavior*, *59*, 65–94.

Perszyk, D. R. & Waxman, S. R. (2018). Linking Language and Cognition in Infancy. *Annual Review of Psychology,* 69, 231-250.

Rhodes, M. & Moty, K. (2020). What is social essentialism and how does it develop? *Advances in Child Development & Behavior,* 59, 1-30.

Rhodes, M., Leslie, S.-J., Bianchi, L., & Chalik, L. (2018). The role of generic language in the early development of social categorization. *Child Development, 89*(1), 148–155.

Rhodes, M., & Mandalaywala, T. M. (2017). The development and developmental consequences of social essentialism. *Wiley Interdisciplinary Reviews: Cognitive Science, 8* (4), e1437.

Rhodes, M., & Gelman, S. A. (2009). A developmental examination of the conceptual structure of animal, artifact, and human social categories across two cultural contexts. *Cognitive Psychology, 59* (3), 244–274.

Scarantino, A. (2010). Insights and Blindspots of the Cognitivist Theory of Emotions. *The British Journal for the Philosophy of Science*, 61(4), 729–768.

Shargel, D. & Prinz, J. (2018). "An Enactivist Theory of Emotional Content", in Hichem Naar and Fabrice Teroni (eds.), *The Ontology of Emotions*, Cambridge: Cambridge University Press, 110–129.

Spelke, E.S. (2017). Core knowledge, language, and number. *Language Learning & Development 13*(2), 147–70

Spelke, E.S., & Kinzler, K.D. (2007). Core knowledge. *Developmental Science 10*(1), 89–96.

Waxman, S. R. (2013). Building a better bridge. In M. Banaji & S. Gelman (Eds.), *Navigating the social world: What infants, children, and other species can teach us.* Cambridge: Oxford University Press.

Waxman, S. R. (2010). Names will never hurt me? Naming and the development of racial and gender categories in preschool-aged children. *European Journal of Social Psychology, 40* (4), 593–610 .

Waxman S. R. & Markow, D.B. (1995). Words as invitations to form categories: Evidence from 12- to 13-month-old infants. *Cognitive Psychology, 29*(3), 257-302.