

Feature article: Emotion & Evolution

Internalized Norms and Intrinsic Motivations: Are Normative Motivations Psychologically Primitive?

[Daniel Kelly](#)¹

[Department of Philosophy](#)
[Purdue University](#)

Introduction

The possibility that normative motivations are basic or psychologically primitive is an intriguing one worthy of more attention. On the one hand, there is a powerful case that human minds are equipped with a psychological system dedicated to norms and norm-guided behavior (Setman & Kelly, forthcoming). On the other hand, there has not yet been a convincing case made that there are any distinct, *sui generis* motivational resources that are unique or exclusive to this system. To the extent that the issue is addressed, many discussions simply proceed as if the motivations that drive different norm-guided behaviors are drawn from a number of different and more basic psychological sources. However, I do not think the possibility that some normative motivations are psychologically primitive has been ruled out.

My modest aim in this piece is to frame and illuminate some of the issues surrounding normative motivation, rather than take a firm position on any of them. I begin by clarifying the key terms in my title of this essay, and unpacking some of the assumptions that underpin its question. I then distinguish four kinds of answers one might give. In this short essay I will not be able to properly develop and evaluate an argument for the view that normative motivations

are psychologically primitive, but I will have some comments about what such an argument might look like, and what it would have to show.

Spotlight on Internalized Norms

Norms are the often informal rules that structure human behavior, regulating what is appropriate, required, prohibited, or permitted. Such rules govern human activities ranging from dress codes and workplace hierarchies to mate selection and courtship traditions; from dining practices and conversational etiquette to family dynamics and religious rituals (Henrich, 2015; Bicchieri, 2016; Gelfand, 2018). Researchers have produced a number of taxonomies that classify norms, some by reference to the behaviors they govern (sartorial norms, dining norms, conversational norms), others by reference to the values they help realize (care norms, purity norms, individualistic norms), and still others by reference to the ways in which they are stabilized (conventional norms, descriptive norms, injunctive norms) (see O'Neill, 2017). Norms can also be distinguished by reference to the functional role they occupy in the psychological economy of an individual person.

Here a key notion is that of an *internalized norm*. The notion has a venerable history in anthropology and sociology, and has recently been taken up by cognitive scientists, evolutionists, and behavioral economists. For example, Gintis (2003) describes internalized norms as “enforced in part by internal sanctions, including shame, guilt and loss of self-esteem, as opposed to purely external sanctions, such as material rewards and punishments” (p. 407). In a recent paper modeling the kinds of evolutionary dynamics that might have produced the capacity to internalize norms, Gavrillets and Richerson (2017) state that “Certain norms are internalized, that is, acting according to a norm becomes an end in itself rather than merely a tool in achieving certain goals or avoiding social sanctions” (p. 1). Finally, Henrich and Ensminger (2014)

about this material, especially Luca Barlassina, Christopher Bennett, Julia Driver, and Stephen Laurence.

¹ Thanks to Michael Brownstein, Peter Carruthers, Taylor Davis, Jeffrey Jensen, Alex Madva, Cain Todd, and Eric Walle for very helpful feedback on earlier versions of this. Thanks also to the participants in the Emotional Norms Conference held at University of Sheffield in July of 2020 for excellent conversations

characterize internalized norms in terms of their “emotional, or motivational, aspect,” claiming that these are marked by the fact that the “desire to adhere to norms and to see them enforced appears to be internally motivated in some fashion. Once internalized, norms become ultimate ends, goals, or values in themselves” (p. 22). While they differ in the specifics, all of these descriptions share the idea that a norm has been internalized by a person when it comes to bear a special kind of connection to her motivational apparatus. Moreover, they all depict that connection as rather direct and robust, suggesting that a person’s impetus to follow a norm she has internalized is insulated from other influences. Relative to rules she has not internalized, her motivation to conform to an internalized norm is less affected by, for example, the presence or absence of material rewards or punishments, the likelihood of social sanction from others, or changes to other relevant aspects of her external environment.

Indeed, the category of an internalized norm is often brought into focus by appeal to other categories of rules whose members are likewise not distinguished by their content, but rather by the different way they are related to an individual’s psychological makeup. Consider an example of what can be called an *incentivized norm*. Imagine someone who loves to drive fast, but stays under a 55 mph speed limit to avoid getting a ticket or losing her driver’s license. If the limit goes up to 80 mph, or if the driver enters a section of the highway she knows is empty of police, she will no longer be motivated to stay below 55, and will indulge her need for speed. Such examples illustrate that people follow some rules merely as a means to something else, some outcome or consequence beyond the rule itself. In more familiar terms, a person is only *instrumentally* motivated to comply with her incentivized norms, and if the external incentives change or disappear, then so too does a person’s proximate motivation to comply with the rule. It is tempting to describe these as cases in which the person is motivated to follow the rule “by” the incentive, but this is slightly misleading. Strictly speaking, the most proximate motivator of the person’s behavior isn’t the external incentive itself, but still some internal psychological state or other. Incentivized norms differ from



Professor Daniel Kelly

internalized norms is that in the former case, whatever the relevant motivating state (a desire, a fear, a goal, etc.), that state is more directly sensitive to the presence or absence of the external incentive than it is to the internally represented rule.

Turning to internalized norms, the idea of internalization can be further developed by appeal to a *norm system*. There is a growing case that human minds have an evolved capacity dedicated to norms and norm guided behavior (Chudek & Henrich, 2011; House et al., 2019; Kelly & Davis, 2018; Kelly & Stich, 2007; Mikhail, 2011; Nichols, 2004, forthcoming; O’Neill & Machery, 2018; Richerson & Boyd, 2005; Sripada & Stich 2007; Tomasello, 2016). While many details remain to be settled, the norm system can be thought of a package of psychological mechanisms that undergird a capacity to “do” norms. This capacity is marked by a person’s propensities to detect and acquire the behavior-guiding rules prevalent in her social environment, to keep her own behavior in conformity with those norms, and to enforce norms by sanctioning those who violate them. Evidence suggests that the psychological machinery dedicated to performing these tasks of acquisition, compliance, and enforcement

exhibits many of the characteristics of so-called System 1 cognition, especially domain specificity and automaticity (Kahneman, 2011). Together these allow the norm system to operate alongside of and with some degree of independence from other psychological processes, including practical reasoning, reflective deliberation, and conscious volition.

Putting this thumbnail sketch of a norm system together with the characterizations of “internalized norms” offered above suggests a straightforward interpretation: a person has internalized a rule once it has come to be represented in her norm system, once it is acquired and stored in this specific part of her mind. And in virtue of coming to occupy the particular functional role carved out by the norm system, the rule also thereby gets connected up to her motivational apparatus in a special and direct way, such that her compliance and enforcement of the norm is different from and not merely, in the sense spelled out above, instrumental (for example, see Fehr & Falk, 2002; Fehr & Gächter, 2002; Gneezy & Rustichini, 2000). Thus, a person has internalized those rules that are represented in her norm system, and those rules are imputed with what can be called *intrinsic normative motivation*.² Unlike incentivized norms, norms that have been internalized activate their associated behavioral tendencies directly, bypassing—and thus independently of—practical reasoning. Indeed, since the influence of a person’s internalized norms on her behavior is typically not mediated by practical reasoning or

reflective decision making, that influence will often have to be inhibited if she is to refrain from acting on the rule.³

One last note: the categories of incentivized and internalized norms do not exhaust the ways that rules can be represented in the human mind or connected to motivation and behavior. People can *merely cognize* rules. They can know about rules without being in any way motivated to enforce or keep their own behavior in line them. For example, a fan may have an exhaustive knowledge of the NBA rule book even though he never plays basketball himself, or a scholar of Ancient Greece may have spent a lifetime developing an encyclopedic knowledge of the norms that governed social life in the heyday of Sparta. Despite their expertise, in neither case are the respective rules internalized in the sense at issue here. People can also create and adopt their own rules. *Avowed norms* are those that an individual voluntarily imposes on himself, typically after explicitly formulating, reflecting on, and deliberately endorsing them. When a person, say, chooses to stop drinking, or decides to write in a journal for 30 minutes every day, he adopts a rule he will try to satisfy and stay committed to going forward. This capacity to choose, endorse, and follow our own rules is crucial for self-determination and identity formation, and is of great interest to philosophers concerned with agency and autonomy (Callard, 2018; Ismael, 2016; Korsgaard, 2009).⁴ There is also reason to think that avowed norms are undergirded by psychological machinery and

² Though it is not an uncommon locution, it is surprisingly tricky to say exactly what might be meant by calling motivation *intrinsic* other than: “not instrumental”. The terms “intrinsic” and “extrinsic” are obviously not specific to motivation, emotion, or even psychology; like “internal” and “external”, or “endogenous” or “exogenous”, they merely mark their subject matter as falling on one side or another of some boundary. As noted above, there’s an important sense in which all motivating states are “internal” to a person’s mind. Moreover, use of these terms in cognitive science is complicated by the recent debates over active and passive externalism and the rise in visibility of different accounts of embodied, embedded, and extended cognition, all of which have blurred the boundary between what is internal and

external to minds in general. Future research is needed to help clarify the issue.

³ Internalized norms may produce an eccentric phenomenology as well. For instance, Stanford (2018) notes that from a subjective point of view, some norms exhibit a “puzzling combination of objective and subjective elements” (p. 2). Ramstead et al. (2016) develop a notion of “cultural affordance” that provides what appears to be a promising way of capturing how a person’s internalized norms can influence her first-person perspective.

⁴ To put this point more carefully, individuals are at least able to choose *some* of the rules they are bound by; see Witt (2011) and Davidson and Kelly (2018) for discussions that distinguish between norms that an individual embraces voluntarily versus those that she has ascribed to her by other people.

motivational resources that are quite distinct from those associated with the norm system (for longer discussion, see Kelly, forthcoming).

Four Views on the Nature of Normative Motivations

Having clarified my subject matter, my question can be further fleshed out: How do internalized norms reliably produce the compliance and punishment behaviors associated with them? What is the nature and character of intrinsic normative motivations? How are these psychological states related to other, more familiar or better understood motivational states contained in human minds?

The sentimentalist tradition in philosophy has inspired a recently influential line of empirical work in moral psychology that can speak to these questions. This perspective has been fruitful, helping to generate evidence that many important norm-guided behaviors and evaluations are driven by specific emotions like anger, contempt, disgust, or shame (Haidt et al., 2001; Nichols, 2004; Rozin et al., 1999; c.f. Prinz, 2009).⁵ Many versions of this view see these normative behaviors and evaluations as being infused by the particular character of the specific emotion to which they are connected. For example, witnessing the violation of an internalized norm that is connected to anger will motivate a piece of behavior that drives the witness to sanction the transgression, but her sanctioning behavior will also be inflected with many of the characteristics associated with paradigmatic instance of anger in general (approach tendencies, heightened arousal, perhaps the distinctive facial expression). Much of this work also assumes a (common but also contested) view of basic emotions: there are only a handful of them, they are relatively discrete, they are humanly universal and have deep evolutionary roots, and they are themselves psychologically basic and intrinsically motivating (Ekman, 1992; Izard, 2007; Ledoux,

2012; Panksepp & Watt, 2011; Panksepp & Biven, 2012).

This perspective suggests what can be called the *Basic Emotions View*; it is really a family of views, but I'll consolidate for ease of discussion. The view suggests straightforward answers to the questions about normative motivations. It acknowledges that not all of the rules a person cognizes are connected to basic emotions, but those that are typically inherit many features of the emotion to which they are connected, including, most importantly, the emotion's intrinsic motivating force.⁶ Translating this into the terminology developed in the last section implies that internalized norms are those that are intrinsically motivating, and an intrinsically motivated norm gets its motivational force from the basic emotion to which it is connected. The Basic Emotions View also suggests another key task performed by the norm system, namely that of executing a *bundling function* of pairing up each rule it acquires with (on this view) one of the basic emotions. The norm system thus draws on motivational resources endogenous to and made available by basic emotions, but directs and shapes them in new and important ways as well. The emotion is thus transformed on both ends. Upstream, its appraisal conditions will be modified, as the set of cues that activate the emotion will be expanded to include those specified by the compliance and violation conditions for the norm. Downstream, the psychological combination of a rule with its associated basic emotion will channel the motivational force rooted in the emotion (along with some of its more resilient characteristics) into the specific compliance and punishment behaviors encoded in the rule.

A second kind of view takes this template and liberates it from the emotions. The *Multiple Building Blocks View* holds onto the account of the norm system and its bundling function, but expands the set of motivational building blocks it can recruit. The view can take many forms,

⁵ Much of this work also recognizes different categories of behavior-guiding rules. Moreover, many researchers identify those rules that enjoy some kind of tight connection to basic emotions as *moral norms*; see especially Nichols (2004).

⁶ See Kelly (2011, 2013) for this kind of account of purity norms and the emotion of disgust, worked out in terms of a byproduct hypothesis concerning the emotion and norms it gets co-opted to help motivate. Violators of purity norms are often thought of as not just wrong but tainted and contaminating.

depending on what a theorist includes on the list of resources from which the norm system is able draw. For instance, it might be the case that some internalized norms are paired with and inherit the motivational force not of basic emotions but of intrinsically motivating mental states like *desires*, *preferences*, or *attitudes* (Arpaly & Schroeder, 2014; Bicchieri, 2016; c.f. Brennan et al., 2013). Alternatively, theorists could appeal to motivational building blocks that take the form of other dedicated psychological systems.⁷ A plausible set of candidates here are mechanisms associated with *social learning*. Evidence suggests that humans are natural social learners (Laland, 2017; Mathew & Perreault, 2015), and are outfitted with dedicated and early emerging psychological machinery that induces young children to spontaneously imitate (and often overimitate) others (Hoehl et al., 2019). This suite of mechanisms appears to be equipped with the kind of intrinsically motivating resources the norm system could easily recruit, especially to ensure people conform to norm-governed behaviors they observe, and thus to comply with the norms that they internalize (Kenward et al., 2011; Kenward, 2012; Keupp et al., 2013). The capacity for spontaneous and intuitive imitation seems to emerge in tandem with core elements of the capacity for norms as well (Schmidt, Butler et al., 2016; Schmidt, Rakoczy et al., 2011; Schmidt & Tomasello, 2012; Vaish et al., 2016). While it seems a relatively easy step to go from behavior imitation to norm compliance, it is less clear how appeal to social learning capacities might account for intrinsic motivations to enforce norms and punish transgressors. However, the Multiple Building Blocks View is amenable to the possibility that for any given internalized norm, the intrinsic motivation to *comply* is supplied by a different psychological building block than the intrinsic motivation to *enforce*.⁸

On any version of these two views, normative motivations will not be primitive. Rather, these motivations will get their identity as *normative* motivations in virtue of the roles they have been recruited to play in the norm system: to ensure reliable compliance with and enforcement of internalized rules. But the intrinsic motivational force that those recruited resources bring to bear on their new functions is antecedent to their being bundled together with their associated norm. This result seems to straightforwardly hold on a third view as well, which can be called the *Basic Affect View*. This kind of picture, inspired by recent work by Lisa Feldman Barrett and her colleagues, suggests that affect itself is a (perhaps the) basic psychological primitive: normative motivations are not primitive states to be sure, but neither are emotions like disgust, anger or shame. None of these are fundamental components of human minds, but are rather constructed out of more basic psychological elements, namely affect and perhaps a propensity to copy others' affective tendencies. (see especially Barrett & Bliss-Moreau, 2009, but also Barrett, 2006a, 2006b, 2017, and Duncan & Barrett, 2007). Application of this perspective to norms and norm-guided behavior is in its infancy (see Theriault et al 2020), but it suggests a promising pathway for future research

The Basic Emotions View, the Multiple Building Blocks View, and the Basic Affect View all depict internalized norms not as being driven by the type of slow, deliberate, effortful cognition associated with reflective reasoning and fully conscious deliberation. Nor, however, do they depict the norm system as an ancient or foundational platform of mammalian or primate minds. Rather, they all portray it as a middle tier kludge⁹, an evolutionary latecomer imposed by a tinkering Mother Nature on a more basic motivational repertoire, the elements of which it harnesses and bundles and channels and thereby

⁷ For example, see Blair's (1995) appeal to a violence inhibition mechanism (VIM); see Nichols (2004, chapter 1) for critical discussion.

⁸ Stich (2019) interprets Henrich (2015) as suggesting something along these lines when he characterizes a proto-norm as "a culturally transmitted (i.e., socially learned) package of psychological states that includes: (i) a desire to engage in a certain pattern of behavior

under specified circumstances; (ii) a desire that other people do the same; and (iii) an emotion elicitor that leads to an agonistic emotion (typically anger or disgust) when one becomes aware that another person is not behaving in the desired way. These emotions can and sometimes do lead to punitive behavior directed at people who do not behave in the desired way" (p. 8).

⁹ See Markus (2009).

transforms. An interesting upshot of these views is that as the norm system recruits more basic motivational states, it does not leave them completely unchanged. Rather, the norm system imposes new tasks on the component parts it recruits, and as those parts are integrated into the functioning of this larger embedding system, their operations are tailored so that they are better able to perform their newly acquired functions. These more basic motivational states often remain recognizable even as they are pressed into novel roles, retaining their identity even as they are repurposed and transfigured by their new circumstances.¹⁰

A fourth and final possibility diverges from this general picture and holds that normative motivations are in fact basic. According to a *Basic Normative Motivations View*, available and endogenous to the norm system is a form of distinctive intrinsic motivation that it can pair with an acquired rule, and that the norm system can access without having to do any recruiting outside of itself. On this view, the norm system does not require any further resources in order to provide intrinsic motivation to an internalized norm.¹¹

Such a view seems clearly coherent. What would it take to show that it is true? What kinds of arguments and evidence could be marshalled in favor of it? This deserves more careful consideration than I can deliver here, but I will end by pointing to a useful template. In his paper “Basic Questions”, Carruthers (2018) argues that curiosity and questioning attitudes are primitive, foundational components of human and animal minds. He construes curiosity as “an affective (desire-like or emotion-like) motivational state whose content is a question” (p. 136); a proponent of the Basic Normative Motivations View might similarly construe an internalized norm as an affective motivational state whose content is an injunction. The proponent could then follow Carruthers’ playbook of trying to establish that

such internalized norms are likewise foundational by showing that the psychological machinery underlying such norms have a deep phylogenetic history.

Carruthers looks to comparative psychology to make the case that curiosity-like states are found in a variety of other animals. This argumentative strategy is worth pursuing for internalized norms as well. It might be more of an uphill battle, however. The most widely accepted view currently seems to be that the full range of social behaviors associated with a psychological capacity dedicated to norms are *not* found among other animals, and that the norm system itself does not have a long evolutionary history, but is indeed uniquely human (for example see Boyd, 2017; Riedl et al., 2012). However, those working on animal cognition have begun pushing back on this (Andrews, 2020; Fitzpatrick, under revision; Vincent et al., 2019; von Rohr et al., 2011).

In addition to these comparative arguments, there are more general evolutionary grounds for taking seriously the possibility that human normative motivation is a relatively recent adaptation not shared with other animals, and is a *sui generis*, psychologically primitive component of our minds. Modern human beings are the product of uniquely powerful forms of gene-culture coevolution and cumulative niche construction (Richerson & Boyd, 2005; Stotz 2010). Normative motivations that are both distinctively human and psychologically basic may have been installed in our minds by the kinds of culture-driven genetic selective pressures that have recently (in evolutionary time) driven our species down its unique evolutionary pathway (Henrich 2015; Sterelny, 2012). Exploring this possibility empirically may have to rely on an argument from exclusion and proceed by a kind of process of elimination, ruling out possible explanations in which the behaviors and intrinsic motivations associated with an internalized norm are accounted for by appeal to some other

¹⁰ See Anderson (2014) for a convincing argument that the kind of “reuse” described here is the rule rather than the exception in the brain, and Richerson and Boyd (2001) for an account of how evolutionary pressures that selected for abilities to coordinate and cooperate at larger scales ended up remodeling human social psychology.

¹¹ Such a view is at least suggested by Sripada and Stich (2007), whose boxological model depicts the compliance and punitive motivations associated with internalized norms as distinct from emotion systems, which are represented by their own, separate box.

identifiable source, be it another psychological system (for imitation, for more sophisticated forms of social learning, for anger or disgust), or some other mental state (a desire, attitude, affect etc.). Once more fine-grained hypotheses about normative motivations have been formulated, emerging brain scan technologies and new experimental techniques will surely help enrich investigations as well.

Conclusion

Even if this last view turns out to be false and normative motivations are not psychologically *primitive*, they can still be psychologically *special* in interesting and important ways. For example, *normative motivation* can still pick a category of mental state that, while not basic, is nevertheless distinctive in the sense that appeals to normative motivations as such—rather than merely to the more primitive building blocks they are constructed out of—may be indispensable for explaining large swaths of human behavior, especially those related to large scale coordination, cooperation, and morality.

I have my suspicions about many of these issues, but no definitive answers or impregnable arguments to offer. I hope to have made some headway clarifying some of the conceptual landscape and drawing attention to the exciting questions that work on norm psychology continues to raise about normative motivation and its connection to affect and emotion. Formulating questions is often a prelude to progress, if not a kind of progress itself. Inevitably questions raise more questions, too. For example, how could we make psychological sense of the idea suggested by some mathematical models that internalization comes in degrees, with “oversocialized” individuals internalizing their norms more fully than those who are “undersocialized” (Gavrilets & Richerson, 2017)? How is the dimension of over- and under-socialization related to the dimension of tightness and looseness recently explored by Gelfand and her colleagues (Gelfand et al., 2011)? How and by what mechanisms do the kinds of norms that a person has already internalized affect their ability to internalize new norms (Hagger et al., 2014)? These too are fascinating questions, and I can’t wait to see our collective attempts to answer them play out.

References

- Anderson, M. L. (2014). *After Phrenology: Neural Reuse and the Interactive Brain*. MIT Press.
- Andrews, K. (2020). Naïve normativity: The social foundation of moral cognition. *Journal of the American Philosophical Foundation*, 6(1), 36-56.
- Arpaly, N., & Schroeder, T. (2014). *In Praise of Desire*. Oxford University Press.
- Barrett, L. (2006a). Valence is a basic building block of emotional life. *Journal of Research in Personality*, 40, 35–55.
- Barrett, L. (2006b). Are emotions natural kinds? *Perspectives on Psychological Science*, 1, 28–58.
- Barrett, L.F., & Bliss-Moreau, E. (2009). Affect as a psychological primitive. *Advances in Experimental Social Psychology*, 41, 167-218.
- Barrett, L. (2017). *How Emotions Are Made: The Secret Life of the Brain*. Mariner Books.
- Bicchieri, C. (2016). *Norms in the Wild: How to Diagnose, Measure, and Change Social Norms*. Cambridge University Press.
- Blair, R. J. R. (1995). A cognitive developmental approach to morality: Investigating the psychopath. *Cognition*, 57, 1-29.
- Boyd, R. (2017). *A Different Kind of Animal: How Culture Transformed Our Species*. Princeton University Press.
- Brennan, G., Eriksson, L., Goodin, R., & Southwood, N. (2013). *Explaining Norms*. Oxford, Oxford University Press.
- Callard, A. (2018). *Aspiration: The Agency of Becoming*. Oxford University Press.
- Carruthers, P. (2018). Basic questions. *Mind & Language*, 33, 130–147.
- Chudek, M., & Henrich, J. (2011). ‘Culture–gene coevolution, norm-psychology and the emergence of human prosociality. *Trends in Cognitive Sciences*, 15(5), 218-226.
- Davidson, L., & Kelly, D. (2018). Minding the gap: Bias, soft structures, and the double life of social norms. *Journal of Applied Philosophy*, 1-21.
- Duncan, S., & Barrett, L. (2007). Affect is a form of cognition: A neurobiological analysis. *Cognition and Emotion*, 21(6), 1184–1211.

- Ekman, P. (1992). An argument for basic emotions. *Cognition and Emotion*, 6, 169-200.
- Fehr, E., & Falk, A. (2002). Psychological foundations of incentives. *European Economic Review*, 46, 687-724.
- Fehr, E., & Gächter, S. (2002). Do incentive contracts undermine voluntary cooperation? *SSRN*.
- Fitzpatrick, S. (Under Revision). Chimpanzee normativity: Evidence and objections.
- Gavrilets, S., & Richerson, P. (2017). Collective action and the evolution of social norm internalization. *Proceedings of the National Academy of Sciences*, 114 (23), 6068-6073.
- Gelfand, M. (2018). *Rule Makers, Rule Breakers: How Tight and Loose Cultures Wire Our World*. Scribner.
- Gelfand, M. et al. (2011). Differences between tight and loose cultures: A 33-nation study. *Science*, 332, 1100 -1104.
- Gintis, H. (2003). The hitchhiker's guide to altruism: Gene-culture coevolution, and the internalization of norms. *Journal of Theoretical Biology*, 220, 407-418.
- Gneezy, U., & Rustichini, A. (2000). A fine is a price. *The Journal of Legal Studies*, 29 (1), 1-17.
- Hagger, M. S., Rentzelas, P., & Chatzisarantis, N. L. (2014). Effects of individualist and collectivist group norms and choice on intrinsic motivation. *Motivation and Emotion*, 38, 215-223.
- Haidt, J. (2001). The emotional dog and its rational tail: A social intuitionist approach to moral judgment. *Psychological Review*, 108, 814-834.
- Henrich, J. (2015). *The Secret of Our Success: How Culture Is Driving Human Evolution, Domesticating Our Species, and Making Us Smarter*. Princeton University Press.
- Henrich, J., & Ensminger, J. (2014). Theoretical foundations: The coevolution of social norms, intrinsic motivation, markets, and the institutions of complex societies. In J. Ensminger & J. Henrich (Eds.), *Experimenting with social norms: Fairness and punishment in cross-cultural perspective* (p. 19-44). Russell Sage Foundation.
- Hoehl, S., Keupp, S., Schleihauf, H., Mcguigan, N., Buttelmann, D., & Whiten, A. (2019). 'Over-imitation': A review and appraisal of a decade of research. *Developmental Review*, 51, 90-108.
- House, B., Kanngiesser, P., Barrett, H. C., Broesch, T., Cebiolglu, S., Crittenden, A. N., Erut, A., Lew-Levy, S., Sebastian-Enesco, C., Smith, A. M., Yilmaz, S., & Silk, J. B. (2019). Universal norm psychology leads to society diversity in prosocial behaviour and development. *Nature Human Behaviour*, 1, 36-44.
- Ismael, J. (2016). *How Physics Makes Us Free*. Oxford University Press.
- Izard, C. (2007). Basic emotions, natural kinds, emotion schemas, and a new paradigm. *Perspectives on Psychological Science*, 2, 260-280.
- Kahneman, D. (2011). *Thinking Fast and Slow*. Farrar, Straus and Giroux.
- Kelly, D. 2011. *Yuck! The Nature and Moral Significance of Disgust*. The MIT Press.
- Kelly, D. (2013). Moral disgust and the tribal instincts hypothesis. In K. Sterelny, R. Joyce, B. Calcott, & B. Fraser (Eds.), *Cooperation and Its Evolution* (pp. 503-524). The MIT Press.
- Kelly, D. (forthcoming). Two ways to adopt a norm: The (moral?) psychology of avowal and internalization. In M. Vargas & J. Doris (Eds.), *The Oxford Handbook of Moral Psychology*.
- Kelly, D., & Davis, T. (2018). Social norms and human normative psychology. *Social Philosophy & Policy*, 35(1), 54 - 76.
- Kelly, D., & Stich, S. (2007). Two theories of the cognitive architecture underlying morality. In P. Carruthers, S. Laurence, & S. Stich (Eds.), *The Innate Mind Vol 3.: Foundations and Future Horizons* (pp. 348-366). Oxford University Press.
- Kenward, B. (2012). Over-imitating preschoolers believe unnecessary actions are normative and enforce their performance by a third party. *Journal of Experimental Child Psychology*, 112, 195-207.
- Kenward, B., Karlsson, M., & Persson, J. (2011). Over-imitation is better explained by norm learning than by distorted causal learning.

- Proceedings of the Royal Society B: Biological Sciences*, 278(1709), 1239–1246.
- Keupp, S., Behne, T., & Rakoczy, H. (2013). Why do children overimitate? Normativity is crucial. *Journal of experimental child psychology*, 116 (2), 392-406.
- Korsgaard, C. (2009). *Self-constitution: Agency, Identity, and Integrity*. Oxford University Press.
- LeDoux, J. (2012). Rethinking the emotional brain. *Neuron*, 73, 653–676.
- Marcus, G. (2007). *Kluge: The Haphazard Construction of the Human Mind*. Houghton Mifflin.
- Mathew, S., & Perreault, C. (2015). Behavioural variation in 172 small-scale societies indicates that social learning is the main mode of human adaptation. *Proceedings of the Royal Society B: Biological Sciences*, 282, 20150061.
- Mikhail, J. (2011). *Elements of Moral Cognition: Rawls' Linguistic Analogy and the Cognitive Science of Moral and Legal Judgment*. Cambridge University Press.
- Nichols, S. (2004). *Sentimental Rules: On the Natural Foundations of Moral Judgment*. Oxford University Press.
- Nichols, S. (forthcoming). *Rational Rules: Towards a Theory of Moral Learning*.
- O'Neill, E. (2017). Kinds of Norms. *Philosophy Compass*, 12(5), 1-15.
- O'Neill, E., & Machery, E. (2018). The normative sense: What is universal? What Varies? In A. Zimmerson, K. Jones, & M. Timmons (Eds.), *The Routledge Handbook of Moral Epistemology*. Routledge Press.
- Panksepp, J. & Watt, D. (2011). What is basic about basic emotions? Lasting lessons from affective neuroscience. *Emotion Review*, 3, 387–396.
- Panksepp, J., & Biven, L. (2012). *The Archaeology of Mind: Neuroevolutionary Origins of Human Emotions*. WW Norton & Company.
- Prinz, J. (2009). *The Emotional Construction of Morals*. Oxford.
- Ramstead, M. J., Veissière, S. P., & Kirmayer, L. J. (2016). Cultural affordances: scaffolding local worlds through shared intentionality and regimes of attention. *Frontiers in Psychology*, 7, 1090.
- Rakoczy, H., and Schmidt, M.F. (2013). The early ontogeny of social norms. *Child Development Perspectives*, 7(1),17-2.
- Richerson, P., & Boyd, R. (2001). The evolution of subjective commitment to groups: A tribal instincts hypothesis. In R. Nesse (Ed.), *The Evolution and the Capacity for Commitment*, (pp. 186-220). Russell Sage.
- Riedl, K., Jensen, K., Call, J., & Tomasello, M. (2012). No third-party punishment in chimpanzees. *Proceedings of the National Academy of Sciences of the United States of America*, 109(37), 14824-9.
- von Rohr, C., Burkart, J., & Van Schaik, C. (2011). Evolutionary precursors of social norms in chimpanzees: a new approach. *Biol Philos* 2011, 26,1-30
- Rozin, P., Lowery, L., Imada, S., & Haidt, J. (1999). The CAD triad hypothesis: A mapping between three moral emotions (contempt, anger, disgust) and three moral codes (community, autonomy, divinity). *Journal of Personality and Social Psychology*, 76(4), 574-586.
- Schmidt, M., Rakoczy, H., & Tomasello, M. (2011). Young children attribute normativity to novel actions without pedagogy or normative language. *Developmental Science*, 14, 530–539.
- Schmidt, M., Butler, L., Heinz, J., & Tomasello, M. (2016). Young children see a single action and infer a social norm: Promiscuous normativity in 3-year-olds. *Psychological Science*, 1-11.
- Schmidt, M., & Tomasello, M. (2012). Young children enforce social norms. *Current Directions in Psychological Science*, 21, 232–236.
- Setman, S., & Kelly, D. (forthcoming). The psychology of normative cognition. *The Stanford Encyclopedia of Philosophy*.
- Sripada, C., & Stich, S. (2007). A framework for the psychology of norms. In P. Carruthers, S. Laurence, & S. Stich (Eds.), *The innate mind: Culture and cognition* (pp. 280-301). Oxford University Press.
- Stanford, K. (2018). The difference between ice cream and Nazis: Moral externalization and the evolution of human cooperation. *Behavioral Brain Sciences*, 41, 1-57.

- Sterelny, K. (2012). *The Evolved Apprentice: How Evolution Made Humans Unique*. The MIT Press.
- Stich, S. (2019) Did religion play a role in the evolution of morality? *Religion, Brain & Behavior*, 1-11.
- Stotz, K. (2010). Human nature and cognitive–developmental niche construction. *Phenomenology and the Cognitive Sciences*, 9, 483–501.
- Theriault, J. E., Young, L., & Barrett, L. F. (2020). The sense of should: A biologically-based model of social pressure. *Physics of Life Reviews*.
- Tomasello, M. (2016). *A Natural History of Human Morality*. Harvard University Press.
- Vaish, A., Herrmann, E., Markmann, C., & Tomasello, M. (2016). Preschoolers value those who sanction non-cooperators. *Cognition*, 153, 43-51.
- Vincent, S., Ring, R., & Andrews, K. (2018). Normative practices of other animals. In A. Zimmerman, K. Jones, & M. Timmons (Eds.) *The Routledge Handbook of Moral Epistemology* (pp. 57-83). Routledge.