Available online at www.sciencedirect.com

ScienceDirect

Current Opinion in Psychology

The neuroscience of moral cognition: from dual processes to dynamic systems

Jay J Van Bavel, Oriel FeldmanHall and Peter Mende-Siedlecki

Prominent theories of morality have integrated philosophy with psychology and biology. Although this approach has been highly generative, we argue that it does not fully capture the rich and dynamic nature of moral cognition. We review research from the dual-process tradition, in which moral intuitions are automatically elicited and reasoning is subsequently deployed to correct these initial intuitions. We then describe how the computations underlying moral cognition are diverse and widely distributed throughout the brain. Finally, we illustrate how social context modulates these computations, recruiting different systems for real (vs. hypothetical) moral judgments, examining the dynamic process by which moral judgments are updated. In sum, we advocate for a shift from dual-process to dynamic system models of moral cognition.

Address
New York University, United States

Corresponding author: Van Bavel, Jay J (jay.vanbavel@nyu.edu, jay.vanbavel@gmail.com)

Current Opinion in Psychology 2015, **6**:167–172

This review comes from a themed issue on **Morality and ethics**

Edited by **Francesca Gino** and **Shaul Shalvi**

For a complete overview see the Issue and the Editorial

Available online 19th August 2015

http://dx.doi.org/10.1016/j.copsyc.2015.08.009

In June 2003, a runaway freight train sped toward downtown Los Angeles at 70 miles per hour [1]. Railway officials estimated that if the train hit downtown LA, it would kill dozens. However, diverting the train to a less populous, blue-collar neighborhood could potentially reduce the cost of damage and the death toll. After deliberating, the officials decided to divert the train, which crashed into several homes, injuring a dozen residents, including several small children. Every day, medical researchers, insurance adjusters, and automobile manufacturers are forced to address a similar moral dilemma: how do we evaluate the costs of human life against other considerations?

Moral questions like these have long been debated within the philosophical community, often through the use of thought experiments designed to probe intuitions of right and wrong. This approach has gained considerable traction in experimental research, where thought experiments like the Trolley dilemma — in which people must decide to kill one person to save five — have become some of the most popular methods in psychology. These ubiquitous moral dilemmas illustrate that changing one simple factor (e.g., pushing a man vs pulling a lever) can profoundly alter moral judgment [2], and offer evidence that emotion and reason can act as dueling psychological inputs. Taken together, these results have provided the empirical foundation for dual-process models of morality (e.g. [3,4]). Yet these hypothetical moral dilemmas fail to capture many of the psychological and neural processes that underpin real moral behavior.

To better understand the mental computations underlying moral judgment and decision-making, recent theories of morality have begun to integrate ancient philosophical debates with rich psychological and biological models. We review research from the dual-process tradition, in which intuitions are automatically elicited and reasoning is subsequently deployed to rationalize or correct for these initial intuitions. We next describe how the mental computations underlying moral cognition are diverse (including personal goals and identities, representations of others' mental states and character, and social norms) and widely distributed throughout the brain. These findings illustrate why models of moral cognition require more than two processes. We also describe how social context modulates these computations, recruiting different systems for real (vs. hypothetical) moral judgments. Finally, we examine the dynamic process by which moral judgments are updated based on new information. Taken together, we argue for a shift away from dual-process models in favor of a more dynamic model of moral cognition.

Dual-process morality

Over the past fifteen years, the field of moral cognition has focused heavily on the roles of intuition and reason in moral judgment. For example, the Social Intuitionist Model [5] states that these intuitions provide an automatic, internal signal that guides moral judgment. According to this model, the human capacity for reason is largely relegated to the role of *post hoc* rationalizing, merely serving to justify initial moral intuitions. Indeed, the human brain responds to harmful acts within a hundred milliseconds — suggesting that moral intuitions emerge very rapidly [6•]. Likewise, the physical act of harming someone feels intuitively wrong to many people,

even if it might save the lives of many others. Accordingly, people are less willing to sacrifice one life to save several others when they have to physically harm someone than when they are emotionally distant from the situation [7]. Early neuroimaging work suggested that these more 'personal' actions evoke a strong response in brain regions associated with affective valuation — like the ventromedial prefrontal cortex (vmPFC) — dovetailing with the theory that automatic intuitions can overpower more deliberate, utilitarian calculations [2]. Similarly, damage to the vmPFC is associated with increased support for a more rational, utilitarian response (e.g. [8,9]). These studies have firmly established the role of intuition in moral cognition.

It remains a source of contention, however, whether or not people can effectively regulate these intuitions. According to many dual-process models, corrective control plays a major role in moral judgment [3,4]. People can override their initial aversion to harming someone in order to save several others, and this utilitarian decision is associated with activation in several brain regions implicated in cognitive control — such as the anterior cingulate cortex (ACC) and lateral prefrontal cortex [10]. This is especially true during high-conflict dilemmas, when strong emotional intuitions clash with utilitarian computations — such as pushing a man off a bridge or smothering a baby to save several other lives. Moreover, cognitive load [11] and temporary disruption of the lateral prefrontal cortex [12,13] impair the ability to enact these utilitarian decisions. Consistent with dual-process models, intuition and reason often work in opposition to generate moral judgments.

Dual-process models: a false dichotomy?

Debates about the precise interplay between intuition and reason have dominated the study of morality. However, developments in cognitive neuroscience lay bare a major problem underlying this debate: there are more than two processes guiding moral judgment (Figure 2; see [14,15]). For instance, one influential review of the neural systems underlying dual process models of moral cognition actually described at least eight distinct brain regions, each of which implements a discrete cognitive process [16]. The complexity of moral cognition is compounded by the temporal dynamics of moral judgment, which unfold over time courses ranging from milliseconds to days. For example, initial reactions regarding harm can be modulated depending on the demands of the situation and as information about the target and normative context is revealed [6•]. These rapidly emerging, highly dynamic cognitive processes involve more than simple corrective control and are critical to understanding moral judgment.

We are not the first to suggest that morality is better characterized as a dynamic system [17,18], nor is this feature specific to morality: the evaluative system is comprised of widely distributed component processes that are highly interactive and sensitive to frames and mindsets ([19]; see also [20]). The field of neuroscience underscores the need to move beyond the simple dual-process dichotomy and embrace models of moral cognition that capture the rich, dynamic nature of human psychology and neuroscience. It would be a lost opportunity if we reduced the rich neuroscience findings into pre-existing, over-simplified psychological models. Instead, we argue that psychological models can benefit from incorporating the complexity and biological realism from neuroscience research. In the next section, we specify several component processes involved in moral cognition and begin to articulate the complex neural, social, and temporal dynamics that underlie moral judgment.

Toward a psychological and neural model

At its core, moral judgment juxtaposes the interests of self and other, such as choosing self-enhancement at another's expense [21•]. The moral self-comprises moral 'oughts' and ideals [22], as well as situational and relational needs [23]: What kind of moral person do you aim to be? Who is a member of your moral community? Which moral codes do you endorse? An individual must continually integrate their own needs with the needs of others, within the broader context of what is socially and morally appropriate (i.e. social norms; [24,25]). Furthermore, representations of other individuals often incorporate mental state attributions (what are the intentions behind this person's behavior? [26]), social categories (is this person a friend or foe? [27]), and reputational information (has this person harmed others in the past? [28]). Thus, the moral value of a given action — its rightness or wrongness — emerges from the integration of a combination of self and other-related information.

This interaction between self and other serves as the foundation upon which other social, societal, and situational pressures are dynamically exerted (see Figure 1). For example, there are times when the needs of the self-outweigh social norms or the needs of others, such as when people steal and harm others for material gain or power or when the others are members of a despised outgroup [21•,29•]. Moreover, the perception of moral ideals or oughts can change depending on the social context in which a moral dilemma is presented [30]. Context can also shift the weighting given to self and other-oriented concerns when computing the moral value of a particular choice or behavior [31–33].

The hypothetical scenarios used to study moral decision-making typically ignore the influence of socio-emotional factors [34] and contextual tensions endemic to real moral dilemmas [21•,35,36]. As such, an individual's response to these dilemmas often fails to reflect how they would truly behave in a more enriched context. When a decision is hypothetical, individuals routinely report that they would