





ARE THERE “MORAL” JUDGMENTS?

David Sackris¹ and Rasmus Rosenberg Larsen²

¹ Arapahoe Community College

² University of Toronto, Mississauga

Original scientific paper – Received: 09/09/2022 Accepted: 09/02/2023

ABSTRACT

Recent contributions in moral philosophy have raised questions concerning the prevalent assumption that moral judgments are typologically discrete, and thereby distinct from ordinary and/or other types of judgments. This paper adds to this discourse, surveying how attempts at defining what makes moral judgments distinct have serious shortcomings, and it is argued that any typological definition is likely to fail due to certain questionable assumptions about the nature of judgment itself. The paper concludes by raising questions for future investigations into the nature of moral judgment.

Keywords: *metaethics; moral judgment; judgment; ontology.*

Introduction

An enduring assumption in the Western philosophical tradition is that moral judgment—namely, judgments about *rightness* and *wrongness*—constitutes a specific type of judgment, distinct from other judgment types. In other words, philosophers have traditionally assumed that when human beings make ‘moral’ judgments, they are doing something typologically (i.e., categorically) different from when they make judgments about other affairs, such as judging that today is Tuesday, that the moon is full, or that $2+3=5$.

As noted by Walter Sinnott-Armstrong and Thalia Wheatley (2012; 2014), this assumption—that moral judgments form a distinct type of judgment—has been at the center of various conversations in the history of moral philosophy. For example, Immanuel Kant is often interpreted as proposing in the *Groundwork of the Metaphysics of Morals* (1998) that a moral judgment is made by a distinct faculty of the mind, and that moral judgments have distinctive properties that other types of judgments lack: only judgments of morality are categorically binding (see also Stich 2006).¹

Most contemporary philosophers seemingly continue to maintain this assumption, as exemplified by the questions that dominate current debates, such as: are moral judgments made using cognitive or non-cognitive processes? Are moral judgments necessarily motivating? Can moral judgments be true? If moral judgment was not assumed to be a distinctive judgment type, then it is not obvious that these questions would at all be meaningful (for an overview of the sorts of questions focused on in contemporary metaethics, see Sayre-McCord 2014; Smith 1994). Note that we do not usually ask such questions about many other judgments. That is, we do not ask if judgments about the weather necessarily *motivate*, or if such a judgment is true in the *same way* as other judgments. We don’t seem to presuppose that there is something typologically distinctive about many of the judgments we make in our day to day lives. They are just, well, judgments.

While it may be intuitively appealing to assume that moral judgments are typologically distinct, the assumption arguably presents challenges that have been scarcely recognized by contemporary philosophers. One such challenge is rather fundamental, as it has to do with providing a clear and

¹ See Sackris and Larsen (2022) for an overview of contemporary metaethics scholarship and the widely held commitment to the position that moral judgment constitutes a distinctive type. See also the anthology by Decety and Wheatley (2015) for contemporary approaches, and Verplaetse (2009) for a historical overview of the (neuro) science of moral judgment.

unambiguous definition of what makes a judgment ‘moral’. That is, if moral judgment constitutes a distinct type of judgment, then we should be able to state what features distinguish this type of judgment from other types of judgments.² Sinnott-Armstrong and Wheatley, for example, endorse such a definitional approach when they state, “A group of things are unified in the relevant way if and only if they share some feature that enables important universal generalizations about its distinctive properties” (2012, 356). This sort of definitional approach can be written out as the logical formula, *A is a B that Cs*, meaning that we must be able to finish the sentence: *Moral judgments are judgments that ‘C’*, where ‘C’ is filled in by some significant feature that all moral judgments share but that is not possessed by other, non-moral judgments (also known as an *Aristotelian* definition, and this form constitutes the basic logic behind scientific taxonomies; see Arp, Smith, and Spear 2015; Berg 1982; Seppälä, Ruttenberg, and Smith 2017).

In this paper we aim to build on prior arguments that have sought to cast doubt on the idea that moral judgments form or instantiate a unified type, whereby all instances of moral judgments have a shared set of features (i.e., Sinnott-Armstrong and Wheatly 2012; 2014; Stich 2006; Sackris and Larsen 2021). Among these contributions, some have questioned the unity of the category of moral judgement (e.g., Sinnott-Armstrong and Wheatley 2012), but so far none have explicitly argued that the category itself must be rejected. Here we provide and test such an argument, namely: that it is not only difficult, but probably also impossible, to define what makes a judgment of the ‘moral’ type (i.e., defining the ‘C’). We contend that common definitions of moral judgment do not sufficiently distinguish a ‘moral’ judgment from other judgments and/or judgment simpliciter. We advance this proposition by first pointing out that the term ‘judgement’ is itself lacking a proper definition, whereafter we critically survey three possible accounts of ‘moral’ judgment: (1) we consider the common position that moral judgment is distinguished from other judgment types by its content; (2) we consider the possibility that moral judgments are distinguished from other judgments by brain processes; and lastly (3) we consider one of the few explicit and contemporary defenses of the position that moral judgments constitute a distinctive kind (i.e., Kumar 2015; 2016a; 2016b). We show that all three proposals fail to meaningfully differentiate ‘moral’ judgments from other judgment types. Finally, we

² Here we use ‘type’ and ‘kind’ interchangeably. We do not have any philosophical commitments as to what constitutes a ‘natural kind’ (for review, see Brzović 2018), nor do we think such nuances will undermine the main argument articulated in this paper. We use ‘kind’ and ‘type’ to mean a distinctive category that is distinguished by some shared property or feature (or set of shared properties or features).

conclude by raising questions for future investigations into the nature of moral judgment.

Before we proceed with our argument, a point of clarification is necessary. We do not mean to argue that the phrase ‘moral judgment’ has no *meaning* (i.e., that the phrase is semantically empty). To most people, it is undoubtedly meaningful when a person utters “I find smoking immoral” or “It is immoral to be so selfish”. When terms like ‘moral’ and ‘immoral’ are used as adjectives, these can convey all sorts of meanings, for example, that one approves or disapproves of certain behaviors. The argument we promote in this paper aims to raise doubt about the position on which such judgments have a distinct typological *referent*, or whether they refer to something that constitutes a distinctive type (i.e., that it is meaningfully distinguishable from what we might intuitively think of as other judgment types). Indeed, what we aim to reject is the idea that all judgments typically referred to as ‘moral’ necessarily have shared features that allow us to infer that they instantiate a universal sub-category distinct from other judgment types. On our view, when people call something a ‘moral judgment’, there is no clear corresponding referent (such as a specific cognitive type or process). The argument we provide here is an attempt to show that there are good reasons to reject the idea of ‘moral judgment’ as a distinctive type.

1. What is a Judgement?

What we see as a key hindrance in the project of defining ‘moral’ judgment is that it is unclear what exactly is meant by the term ‘judgment’. Ostensibly, any attempt to define ‘moral judgment’ should begin with defining what ‘judgment’ is, and only thereafter define what differentiates a ‘moral’ judgment from ordinary and/or other types of judgment (i.e., as captured in *A is a B that Cs*) (e.g., Arp, Smith, and Spear 2015; Berg 1982; Seppälä et al. 2017). For example, when biologists define and taxonomize newly discovered novel-looking organisms, they must first determine whether it is, say, a vertebrate animal, and only thereafter whether it is a mammal, reptile, bird, etc. If it is a mammal, then it must be determined whether the animal falls under any of the various species-categories in the mammalian taxonomy or whether it can be genuinely distinguished from its evolutionary siblings and ancestors, namely, the ‘C’ in *A is a B that Cs*. Consider, for example, recent controversy concerning whether ‘Brontosaurus’ and ‘Apatosaurus’ constitute two distinctive species of dinosaur. Undoubtedly, extent fossils indicate that the animals in question are part of the diplodocid family, but we only have two distinctive species if we can find some feature that reliably distinguishes one set of bones from another (Osterloff, n.d.).

However, in contemporary moral philosophy, this standard procedure in taxonomic definitions has never really been strictly applied. For example, in Richmond Campbell’s 2007 paper entitled, “What is Moral Judgment?”, there is no discussion of what a ‘judgment’ simpliciter might be (i.e., there is no typological definition of judgment: the ‘B’). Whether all judgments have anything in common, or if there are several different types of judgment in general is not discussed. Campbell does state that “Moral judgments are (or express) states of belief” (2007, 321), but says nothing about what makes a ‘moral’ judgment different from other types of judgments. Indeed, speaking to Campbell’s claim concerning ‘moral’ judgments, we might ask whether or not it is true that *all* judgments express states of belief?³ If so, then his definition fails to differentiate anything distinctive.

Similarly, in a 2015 paper by Victor Kumar entitled, “Moral Judgment is a Natural Kind”, the first line of the paper reads: “Moral judgments seem to be different from other normative judgments, even apart from their characteristic subject matter” (2015, 2887). In this paper, Kumar is distinguishing moral judgment in part by its supposedly distinctive subject matter, but like Campbell, there is no clear attempt at defining ‘judgment’, let alone what a ‘normative’ judgment is (see also May and Kumar 2018).

Further examples come to mind. In Joshua Glasgow’s (2013) exploration of the phenomenological essence of moral judgment, there is no discussion of judgment itself. Even in Michael Smith’s modern classic, *The Moral Problem*, in which he characterizes “the central organizing problem in contemporary metaethics” (1994, 11) as an inability to satisfactorily define moral judgment (given the commitment to a Humean account of motivation), there is little discussion of what ‘judgment’ might be. Smith does hold that moral judgments necessarily motivate, but he does not clarify whether moral judgments are the *only* judgments that necessarily motivate (see especially chapter 3).

In these contemporary publications, it would appear as if the reader is supposed to have a kind of fundamental or intuitive understanding of what a ‘judgment’ is; and thereby also, what a ‘moral’ judgment is. But do they? Apparently, there are good reasons to be skeptical about philosophers intuitively knowing or agreeing on what a ‘judgment’ is, or even supposing that the term ‘judgment’ picks out a single, uniform cognitive process. In

³ Campbell goes on to add that the view that moral judgments are beliefs is in conflict with other beliefs that we might have about moral judgments, as well as beliefs we might have about belief itself, so even if the statement “Moral judgments are (or express) states of belief” was a kind of definition, it wouldn’t be a widely accepted one. Similarly, Smith’s (1994) is focused on this same problem.

Wayne Martin's historical survey of philosophical theories of judgment he concludes:

In large part, the history I recount is a history of philosophical failure. In each study I show how seemingly promising approaches to judgment led more or less directly to theoretical impasse. The problem of judgment, it turns out, proves remarkably resistant to solution—even across a diverse range of disciplines and methodologies. (Martin 2006, 7)

We similarly propose that it would be too hasty to assume that philosophers know what makes a judgment 'moral'. *Prima facie*, what we call 'judgments' likely result from a diverse set of cognitive processes, and those processes may or may not result in end products that have shared features. Therefore, there is legitimate and fundamental reason for suspicion when the task of defining 'judgment' is itself ignored in the literature that aims at defining 'moral judgment'.

What is arguably missing in the field is a clear definition along the following lines: A 'moral' judgment is a 'judgment' that 'Cs'. And it may be speculated that this lack of a clear definition of 'moral judgment' is rooted in the lack of a clear understanding and shared consensus about how to define judgment itself. If judgment is not well understood or if we have unfounded assumptions concerning human judgment, it is going to be that much harder to successfully differentiate 'moral' judgments from other judgment types.

It is to such attempts at defining 'moral' judgment—the 'Cs'—that we now turn.

2. Defining Moral Judgment on the Basis of its Content

In this section, we consider what we shall refer to as the *content approach*, namely, those attempts that aim to define moral judgment by stating what moral judgments are *about*.⁴ That is, when a human being has made a judgment, what makes that judgment 'moral' has to do with the content of said judgment. On this approach, the definitional schema (A is a B that Cs) would be filled out in the following way: Moral judgments are judgments about a certain content.

⁴ The content approach is also found in aesthetics. See for example Carroll (2012; 2015); cf. Sackris and Larsen (2019; 2023).

Several philosophers seem to make just this claim. For instance, Kumar partially endorses such a criterion by referencing a “characteristic subject matter” of moral judgments (2015, 2887). Another contemporary philosopher who seems to venture in this direction is Jesse Prinz, when he defines moral judgments as judgments about those *behaviors* we take to be moral (2007, 47-49). Smith appears to be saying something similar, as he believes platitudes from everyday discourse surrounding morality and moral judgment can be used to define it:

To say that we can analyze moral concepts, like the concept of being right, is to say that we can specify which property the property of being right is by reference to platitudes about rightness. (Smith 1994, 39)⁵

The idea seems to be that we have a commonsense understanding as to what moral judgments are typically *about*.

Although the content approach for differentiating ‘moral’ judgment from other types of ‘judgment’ is certainly appealing, we believe such a definition is both problematic and fundamentally implausible.

First, many philosophers have readily admitted that it is difficult to define what exactly moral philosophy is about (i.e., its content). G.E. Moore famously claimed that ‘good’ itself is undefinable (1903). More recently, James Dreier states “We should just admit that it may be vague whether a given judgment is moral or not” (1996, 411). Owen Flanagan highlights the difficulty of defining (*a priori*) what features of actions make them moral issues (1993, 17). Similarly, Shafer-Landau (2015) also suggests that ‘morality’ is undefinable.⁶ These remarks suggest that moral judgments do *not* have a “characteristic subject matter” if there is well-recognized vagueness as to whether a given judgment counts as a moral one.

⁵ Smith does say “there are limits on the kind of content a set of requirements can have if they are to be moral requirements at all, as opposed to requirements of some other kind” (1994, 40). Of course, Smith doesn’t say what those limits are exactly, and that is just our point. If a person can consider almost any issue a moral one, it is hard to see how far commonplaces about morality will get us to a meaningful distinction between the ‘moral’ and the ‘non-moral’.

⁶ See his “Introduction”. Additional examples can be given. Richardson (2018), in his *Stanford Encyclopedia of Philosophy* entry on moral reasoning states “[W]e will need to have a capacious understanding of what counts as a moral question. For instance, since a prominent position about moral reasoning is that the relevant considerations are not codifiable, we would beg a central question if we here defined ‘morality’ as involving codifiable principles or rules”. Svavarsdottir admits that “it is of course notoriously difficult to say what distinguishes moral judgments from other evaluative or normative judgments” (1999, footnote 6).

Perhaps it is difficult to define morality based on what it is *about* precisely because it appears that almost any behavior can be moralized if persons or a given society decides to moralize it (e.g., Fiske and Rai 2014; Haidt 2012): the problem here is that if any content can be identified as the subject of a moral judgment, then it would be futile to search for content that can serve as a unique identifier. As Sinnott-Armstrong and Wheatly point out:

The mere fact that people tend to group a set together under a name does not show that there is any single feature that enables significant generalizations about all and only things in that group. (Sinnott-Armstrong and Wheatly 2014, 457)

Given the great diversity of content (e.g., behaviors, events, etc.) that have been moralized by human beings (e.g., dietary restrictions, honor killings, sexual relations; see, Fiske and Rai 2014) it is hard to see how studying this hodgepodge could tell us in *what way* moral judgments differ significantly from the judgment of other behaviors and events. If any human activity or event can serve as the basis-content for a moral judgment, then it follows that moral judgments cannot be distinguished from other judgments on the content approach.

Recognizing this line of criticism, a proponent of the content approach might try to refine their view and settle on something like ‘On the surface moral judgments may seem to be about a disunified class of human activities, but actually they are unified by a focus on preventing harm’.⁷ Claims like this, however, force us to look for harms to accompany *every* moral judgment, and it often doesn’t matter if the supposed harms are dubious. Judgments about the wrongness of masturbation, cannibalism, sex outside of wedlock, or homosexual relations do not appear to have anything to do with harm. Finding who is supposed to be harmed by such behaviors requires a good deal of mental gymnastics. Yet, they are all examples of behaviors (i.e., content) that has or currently is being moralized.

The defender of the content approach may also try to fall back on vagueness. For example, there are many vague terms in use (e.g., baldness, heap), which philosophers typically acknowledge as less problematic than they may appear; and perhaps ‘moral judgment’ is just such a vague term. Just as we know a bald person when we see one, we easily recognize core

⁷ See Sinnott-Armstrong and Wheatly (2014, 457-459) for a more thorough discussion of this point, and an able refutation.

cases of moral judgment. That there are borderline or questionable cases doesn't show that we lack a good definition of what is 'moral' content.

To fall back on vagueness seems appealing, but it is unsatisfying for at least one central reason: there is scarce agreement on what counts as 'core' cases of moral content, as persons and cultures consider different and non-overlapping contents as 'core' instances of moral judgment. Even with stereotypical moral cases such as 'murder is immoral' we find plenty of examples where cultures and persons believe that some acts that others would consider 'murder' can be morally justified. For example, genocides have been seen by their perpetrators as morally virtuous acts (e.g., Fiske and Rai 2014). When we use a term such as 'bald' to refer to a person without any hair, no one would disagree that we are using the term appropriately. It is the boundary cases that people disagree about, such as when a person is in the process of losing their hair. Falling back on vagueness is therefore not a convincing defense for the content approach since there is disagreement concerning the 'core' cases.

More significantly, we may legitimately wonder why someone would believe that it is possible to define a judgment type by its content in the first place. To attempt to define a judgment type by what is being judged seems to be the wrong way of approaching things. If 'moral' judgment really is a distinctive type of judgment, it would seem that we would want to identify the distinctive brain or mental process *first*, and then look for what kinds of things trigger *that* or *those* processes. As an analogy, we might be able to catalog items that typically cause people to become sexually aroused, but we would not want to identify or define sexual arousal with those items—since, after all, sexual arousal is a complex neurological and mental process—and we would likely admit that a person could, in theory, become sexually aroused by anything. There may be reliable triggers in large parts of any given population, but sexual arousal is not defined by these triggers.

3. Defining Moral Judgment on The Basis of Its Brain Processes

In this section, we consider whether moral judgments can be distinguished on the basis of typologically distinct cognitive processes. That is, if a judgment is understood as the result of a cognitive *process*, it then seems possible that this process itself could be sufficiently different from other processes when making a *moral* judgment. One way to pursue this strategy, it seems, is to find seemingly uncontroversial issues of *rightness* and *wrongness* (e.g., charity, murder, etc.) and investigate how people judge *those* issues, and whether there are similar or different cognitive processes

undergirding these judgments. If there is, this would perhaps be a good reason to think that moral judgment is a distinctive kind of judgment. If there is not, this would be a reason to doubt that moral judgment is a distinctive judgment type.

In studies conducted in 2001 and 2004, Joshua Greene and others used neuroimaging technology to scan the brains of individuals considering standard moral dilemmas such as the *Trolley problem* and *Footbridge problem* (from Foot 1967; Thomson 1976). Although Greene's stated goal was to debunk the belief that deontological moral judgments were the result of pure reasoning processes (2008, 36), his findings indicate that moral judgments are not reached via a single area/system of the brain or a unified set of brain processes. In Greene's study, when subjects considered dilemmas that directly involved themselves, what Greene refers to as 'personal' moral dilemmas (e.g., would you shove one person from a bridge to stop a trolley from killing five people), areas of the brain that are believed to be involved with emotion-processing were activated. When subjects considered more 'impersonal' moral dilemmas (e.g., flipping a switch from a distance to divert a trolley from killing five and as a result killing one person instead), their judgments appeared to be correlated with activity in areas of the brain responsible for (non-emotional) cognitive processes:

Contemplation of personal dilemmas produced relatively greater activity in three emotion-related areas: the posterior cingulate cortex, the medial prefrontal cortex, and the amygdala [...]. At the same time, contemplation of impersonal moral dilemmas produced relatively greater neural activity in two classically "cognitive" areas of the brain, the dorsolateral prefrontal cortex and inferior parietal lobe. (Greene 2008, 43-44)

Greene did not initially take his findings to show that moral judgment cannot be meaningfully classified as a distinct category of judgment; it is only later that he has come to embrace such a conclusion, stating: "I believe that moral judgment is not a natural kind at the cognitive level" (2015b, 40).

Although it is excruciatingly difficult to identify functional areas of the brain with high confidence (e.g., Marek et al. 2022), what is important for our purposes is that Greene et al. (2001; 2004) showed that different areas of the brain were functioning when confronting different kinds of moral scenarios and/or problems. This reliance on different areas of the brain for forming judgments about what many would readily classify as 'moral'

dilemmas should give us pause. If different functional areas of the brain are being relied on to make judgments in these cases, then this should at least be reason enough to begin questioning our confidence in the assumption that there is a *single* brain process or system that can be identified with the term ‘moral judgment’. In fact, the opposite appears to be true if we follow Greene’s findings: different areas of the brain will be called on to form a judgment depending on the context. If there is no single process (or pattern of processes) that can be identified when test subjects make judgments that are assumed to be moral, then brain processes cannot serve as the basis for typologically differentiating moral judgments from other judgment types. Furthermore, if different brain processes are active during the ‘moral’ judgment process, this might give us reason to think that the resulting judgments from these different areas of the brain might well have different properties (e.g., beliefs in some cases, desires in others).

Building on Greene, Fiery Cushman and Liane Young (2011) have found that patterns of moral judgment can be attributed in part to regions of the brain responsible for the attribution of intentions and causation, general reasoning process that might be engaged in a variety of judgment types. From this they conclude that our moral judgments are “derived” from more general judgment forming processes (2011, 1053). Their findings suggest that there is no distinct ‘moral’ judgment type, but that instead the brain employs a general reasoning process that it applies to a diversity of issues. When reviewing the evidence considered in both Greene (2008) and Cushman and Young (2011), Borg and colleagues (2011) reached a similar conclusion:

Consistent with their structure in the deep brain and given their participation in negative judgments in many contexts, the role of the anterior insula and basal ganglia in judging an act to be morally wrong likely represents a general role for these regions in encoding negative valence and avoiding aversive stimuli rather than a unique role in contributing to negative moral verdicts. (Borg et al. 2011, 408)

Across the board, these researchers are attributing the process of moral judgment formation to parts of the brain that play more general judgment formation roles. In short, these brain processes are not exclusively reserved for what we call ‘moral’ judgments.

Perhaps more significantly, a neuroimaging study by Borg and colleagues (2011) found evidence that the brain may rely on separate neural systems for reaching positive and negative ‘moral’ verdicts (2011, 409). If there are separate systems for reaching positive and negative ‘moral’ judgments,

then how do we reconcile this with the assumption that moral judgments are the result of a unified process about which we can make meaningful generalizations?

There is additional evidence for the disunity of ‘moral’ judgment at the level of brain processes. Liane Young and James Dungan conducted a review of neuroimaging research on moral judgment and concluded that “morality [relies] on domain general-processes which are housed in many parts of the brain (...) morality is virtually everywhere in the brain” (2012, 1).

Again, this conclusion has been reached by other researchers, such as John Decety and Jason Cowell, when they write:

What has become clear from social and clinical neuroscience research is that there is no unique center in the brain for moral judgment. Rather there are interconnected systems that are not domain specific but support more domain-general processing, such as affective arousal, attention, intention, understanding, and decision making. (Decety and Cowell 2014, 528-529)

At this point, the cognitive science community has coalesced around the following conclusion: there is no single area/system of the brain, nor any stable and clear process patterns, that play the primary judgment formation role when people make judgments about what we typically classify as ‘moral’ issues. They now believe that the formation of moral judgments is dependent on a variety of areas of the brain, and these areas are not primarily devoted to moral judgment; that is, what we call ‘moral’ judgment is derived from functional areas of the brain primarily devoted to other tasks.⁸

This seems to imply, then, that the process of making a moral judgment likely shares features with what we previously would have taken to be other, distinct judgment types. Thus, if moral judgments are derived from functional areas of the brain that are primarily devoted to other tasks, then it is unclear how this can serve to typologically differentiate ‘moral’ judgments from judgment simpliciter or other types or judgments.

If we step back a moment and consider this approach to understanding moral judgment, we can see that it was likely to yield such unsatisfying

⁸ See also Sackris and Larsen (2022) and Sackris (2022) for overviews of neuroscientific research on moral judgment formation. They reach a similar conclusion to the one formed here: at this time, moral judgment formation has not been reliably correlated with any distinctive brain process, set of processes, or set of brain areas.

results. As we stated at the beginning of this paper, if we were to ask ourselves whether we think that *judgments* are made by a single area of the brain or whether *judgment* called on the same set of processes in every context, our *prima facie* answer would most likely be ‘No’. The term ‘judgment’ is used to capture the result of an array of psychological processes: from almost instantaneous belief formation (e.g., when a person judges that another person is attractive) to the result of a careful deliberation process (when a person offers the solution to a complex problem in mathematics). Surely these disparate psychological processes must correspond to different brain processes or systems. If it seems unlikely that ‘judgment’ itself could be successfully identified with a distinct set of brain processes/single region of the brain using neuro-imaging technology, then given the variety of contexts that give rise to ‘moral’ judgments, we also have little reason to think that ‘moral’ judgment could be successfully identified with a single distinctive brain area or set of processes that are easily distinguished from other judgment types.

4. Defining Moral Judgments on the Basis of Conceptual and Introspective Features

In this section, we consider whether moral judgments can be typologically defined through conceptual and introspective analysis. That is, even if moral judgment is not unified by its *content* or at the level of *brain processes*, perhaps the way we conceive of and conceptualize moral judgments unifies them as a distinctive type or kind. Afterall, diverse processes could still yield similar results in specific, important respects. For example, we may always conceive of moral judgments as having certain, necessary features that other judgment types simply do not have. On such a view, it is not the content being judged, nor the brain processes engaged that make a judgment a ‘moral’ one, but instead the way the judger reflectively thinks about their judgment that differentiates it from other judgment types.

Such a view was put forward by R.M. Hare (1981, 53-56), as he defined ‘moral judgments’ as judgments that are conceived as: universalizable, prescriptive, and overriding (for a criticism, see Sinnott-Armstrong and Wheatley 2012). According to Hare, the key feature of the term ‘moral’ is its ‘overriding-ness’. We make a great many prescriptions in our ordinary lives; that is, there are various uses of the word ‘ought’ and ‘must’. However, when two prescriptions conflict, the one that is conceived as being ‘moral’ will take precedence over the other prescription (1981, 55-56).

A variant of Hare’s approach has gained some appeal among contemporary philosophers, where it has recently been advanced by Victor Kumar in a series of papers (2015; 2016a; 2016b). Adding to Hare’s definition, Kumar claims that speakers conceptualize moral judgments as (1) serious, (2) general, (3) authority-independent, and (4) objective (e.g., Kumar 2015). By “serious”, Kumar means, like Hare, that moral judgments are “overriding” in relation to other judgments. By “general”, Kumar means that moral judgments are conceived of as not bound by place and time. By “authority-independent”, Kumar means that people conceive of moral/immoral acts as being right/wrong even in circumstances when some authority says that the act is permissible/impermissible to commit. By “objective”, Kumar means that moral judgments are conceived as akin to judgments about matters of fact. Overall, Kumar argues for the position that moral judgments constitute a distinctive type of judgment (he argues that they are a natural kind) based on his claim that they share this cluster of four features that distinguish them from other judgment types.

As Kumar is one of the few contemporary authors who explicitly follows Hare’s approach and argues for the view that moral judgments constitute a distinctive kind or type, we will carefully consider his position.

On Kumar’s view, the four features outlined above form a “homeostatic property cluster” and he proposes that “the human cognitive system is organized in such a way that the four features have a nomological tendency to cluster together” (2015, 2896). That is, these four features co-occur in a law-like manner and distinguish moral judgments from other kinds of judgments (2015, 2889-2890). Kumar maintains that these shared features indicate that moral judgment is, in fact, a unified phenomenon and the presence of these four features is what distinguishes moral judgments from other judgment types.

We might ask why Kumar settles on exactly these four features (and thereby neglects others)? According to Kumar, the four features can be derived from analyzing research on the so-called *moral/conventional* distinction. This body of research has been interpreted to show that individuals (including children as young as three years old) can reliably distinguish between two fundamentally different forms of (normative) violations, namely, moral and conventional violations (cf., Gilligan 2016; Witherell and Edwards 1991). As reported in a handful of studies, when participants are asked to explain the basis for this distinction, some will readily classify moral violations as more wrong than conventional violations; they see moral violations as more authority-independent; they see moral violations as time and place independent; and they explain the wrongness of moral violations in terms of the harm they cause to others,

which they do not typically do for conventional violations (e.g., Nucci and Turiel 1978; Smetana 1981; Turiel 1983; for a recent review see Margoni and Surian 2020).

It is this routine distinction between conventional violations and moral violations, and the rationales that study participants typically give for such a distinction, that leads Kumar to conclude that it is features (1)-(4) that mark out moral judgments as distinctive.

Although Kumar’s approach appears to be based on careful, empirical observations, it nevertheless faces the problem of being open to clear counterexamples. Indeed, for Kumar’s definition to fail, all that needs to be shown is that there are some clear cases of ‘non-moral’ judgments that have the four properties Kumar has identified; or that there are judgments typically classified as moral that lack one (or more) of the four features identified by Kumar.

Evidence of such counterexamples are abundant in the literature. For instance, a study by Daniel Kelly and colleagues (2007) shows that subjects are reliably willing to view their moral judgments as non-objective when considering certain moral vignettes. Similarly, survey-studies by James Beebe and David Sackris (2016) and Geoffrey Goodwin and John Darley (2008) have shown that people do not necessarily conceive of morality as objective.⁹ To be more specific, empirical work has consistently demonstrated that individuals do not always conceive of moral judgments as objective or generalizable. Therefore, Kumar’s claim that moral judgments share the universal feature of objectivity can be refuted on empirical grounds.

When confronted with such strong evidence, Kumar does not disregard the validity of the data, but states instead that the “findings may show that morality is not conceived as universal, but they do not show that morality is conceived as no more general than convention” (2015, 2899). Kumar continues this line of defense when he goes on to say, “a few deliberately chosen cases in which the components come apart is quite a long way from disconfirming evidence” (2015, 2899). Finally, Kumar also states that his theory “cannot be refuted simply by pointing to cases of moral judgment in which *one* of the features are absent” (2015, 2901, our italics).

⁹ Kumar admits in footnote 5 of his (2015) that Goodwin and Darley have shown that subjects do not treat all moral judgments as objective, however he doesn’t take this as evidence against his view, as he attributes it to study participants doubting that such disputes could be rationally adjudicated. He does not discuss Beebe and Sackris. A recent paper by Paul Rehren and Sinnott-Armstrong (2022) indicates that individuals are quite willing to change their mind regarding earlier moral judgments, which also might lead us to question attributions of objectivity.

What, then, would count as disconfirming/falsifying evidence of Kumar's definition of moral judgment? He says that "if clearer counterexamples could be produced, in which judgments that are intuitively classified as moral lack several of the four features, that would count against the view" (2015, 2903). That is, Kumar seems to believe that finding a judgment that is clearly conceived of as moral, yet lacks two of the four features, is the only legitimate way to refute his view.

Although we already see his view as defeated by the empirical evidence discussed and we fail to see why anyone should accept such a high bar for the refutation of Kumar's definition that he sets, we think that it is indeed possible to accommodate Kumar's challenge. Admittedly, it is difficult to conceive of a paradigm moral judgment that is widely taken to lack two of the four identified features. Part of the reason for this is, as we discussed above, that there is general disagreement as to what counts as a moral judgment. However, we will give it a try. Here is what we consider to be a moral judgment that is both *not* serious and *not* authority-independent:

You shouldn't smoke

If we tell someone that they shouldn't smoke, unless we are ourselves medical doctors or spent our time reading medical journals, our judgment is almost entirely authority *dependent*; in fact, someone's advice against smoking might be based entirely on the Surgeon General's warning placed prominently on the pack of cigarettes. Such a person might also be willing to change their judgment if the Surgeon General changed their advice. Second, many people evidently do not take such judgments to be all that serious, that is, these judgments don't necessarily override all other interests, such as their own or the person they might tell this to: most people likely won't knock the cigarette out of someone else's mouth or crush up their entire pack, and probably doesn't expect the person in question to quit that very instant.

Although smoking primarily effects the person who engages in it, it does affect others via second-hand smoke and by setting a bad example, say, to impressionable children. However, even though said behavior has the potential to negatively impact others, we do not actively view such behavior as a *serious* matter that must be *immediately* attended to. For example, a smoker might say to themselves "I really need to stop smoking in front of my children" but then later in the day go ahead and do so anyways for a variety of reasons.¹⁰

¹⁰ 'You should drive an electric car' might be another example of a moral prescription that is both not serious and authority-dependent.

If the smoking-example is unconvincing to some readers, they might consider the following religious injunctions to be examples of moral judgments that are based on authority and not considered all that serious or over-riding by the practitioners. For example, a practitioner of Mormonism is not supposed to consume “hot drinks” on the authority of Joseph Smith but may do so on occasion as they don’t view the injunction as all that serious. On further consideration, one might claim that many human beings base some of their moral judgments on some version of a religious framework, which seems to imply that such judgments by definition are (1) authority-dependent (e.g., the written dogmas), (2) non-objective (e.g., amenable by higher institutional authority), (3) non-general (e.g., rules only apply to the religious cohort), and (4) non-serious (e.g., dogmas will not necessarily overrule other judgments).

We may also refute Hare and Kumar’s approach by identifying judgments that are not typically conceived of as moral yet have the properties of being (1) serious, (2) general, (3) authority independent, and (4) objective. That is, there are seemingly clear examples of non-moral judgments that have all four features identified by Kumar.

For example, consider a philosopher who, after reviewing a great number of arguments, makes the judgment that moral sentimentalism is true—and not moral cognitivism—such a judgment is perhaps regarded by said philosopher as serious in that it overrides some of his earlier judgments about the nature of morality, general, authority-independent, and objective. Does this make it a moral judgment?

This example demonstrates what we see as perhaps the biggest challenge for Kumar’s definition, namely, that it is too *inclusive*. Indeed, there is an endless array of judgments that have all four features that Kumar outlines, yet are by no means traditionally considered to be moral judgments. We imagine that some art critics take their judgments to be serious, general, authority-independent, and objective. We suppose that most scientific judgments have these properties as well. When Galileo judged that the Earth circum-navigated the sun, this was serious (it overrode other judgments), general, authority-independent, and objective. Are all these judgments also moral judgments? On Kumar’s view, they most certainly are. And perhaps for that reason alone, Kumar’s definition falls apart.

While Kumar’s definition of moral judgment seems to admit of counterexamples (e.g., as being empirically falsified or too inclusive), we could also raise a concern about the evidence it rests upon. First, the significance of the moral-conventional distinction has for many years been disputed (e.g., Witherell and Edwards 1991). For example, some

researchers have demonstrated that the way a question is framed has a significant influence as to whether study participants construe the violation of a norm as moral or conventional (e.g., Margoni and Surian 2020).

It may be that in some cases moral judgments are conceived as more serious, or more objective than other judgments; however, in other contexts it may be that moral judgments are not considered all that serious or taken to transcend a given culture. The results from survey-studies (e.g., Beebe and Sackris, 2016; Goodwin and Darley, 2008) indicate that the relationship between moral and conventional violations is much more akin to two different end points on a *single* scale instead of two distinct *kinds* of normative violations. If we asked four-year-old children whether they should ever touch a hot stove, they would likely say “No” regardless of whether an authority figure told them it was okay to do so or even if they were in a foreign setting. On Kumar’s account, this would make prohibitions against touching hot stoves seem to fall into the category of moral injunctions (especially since on his own account not all four features have to be present). If this seems like an inappropriate categorization, then his definition of “moral judgment” has likely missed its mark.

5. Concluding Remarks

In this paper we have argued that the long-standing philosophical assumption that moral judgments are typologically distinct from other judgment types should be questioned, and possibly rejected. In an Aristotelian framework where definitions adhere to the logical constraints —A is a B that Cs—philosophers and scientists have yet to offer clear definitions of the ‘Cs’. Further supporting the conclusion that moral judgments are not typologically discrete is the observation that philosophers often omit defining what a *judgment* is. Logically, it is impossible to build a sufficiently meaningful taxonomy without defining the genus (i.e., the ‘B’) that we are attempting to place species under (i.e., the ‘A’ that ‘Cs’).

As contended throughout this manuscript, although philosophers rarely say this explicitly, they seem to sometimes suppose that moral judgments are always arrived at via some single uniform process. From such a position, it is intuitive to stipulate that there is a ‘moral judgment’ area in the brain, or that moral judgments must *always* have certain processes or features. However, neuroscientific evidence tells a different story, that judgments generally are formed in a variety of ways. The same seems to be true of moral judgments. Some moral judgments are arrived at almost instantaneously; other moral judgments are only arrived at as a result of

deep reflection. Admittedly, the products of different processes can still have similar properties. However, if we recognize that moral judgments can be arrived at in these seemingly very different ways, why are we so tempted to imagine that they must have shared, necessary features?

If we can use the word ‘judgment’ to describe the result of highly diverse cognitive processes that range from instantaneous belief formation to the production of a proof in propositional logic, we have to consider the possibility that what we call ‘moral judgments’ can typically be the result of highly diverse cognitive processes as well; we should also entertain the possibility that these processes are capable of yielding differing results: perhaps in some cases beliefs, in some cases desires, in some cases some third mental state, such as *besires*.¹¹ This doesn’t mean we should throw out the term ‘moral judgment’; but it does mean that we may need to reconceive our inquiry into judgments that we typically refer to as ‘moral’ and recognize that few sweeping generalizations about such judgments are likely to be true.

Acknowledgments

Many thanks to the anonymous reviewers for this journal.

REFERENCES

- Altham, James. 1986. “The Legacy of Emotivism.” In *Facts, Science, and Morality: Essays on A.J. Ayer’s Language, Truth and Logic*, edited by Graham McDonald and Crispin Wright, 275-88. Oxford: Blackwell.
- Arp, Robert, Barry Smith, and Andrew Spear. 2015. *Building Ontologies with Basic Formal Ontology*. Cambridge, MA: MIT Press.
- Berg, Jan. 1982. “Aristotle’s Theory of Definition.” *Atti del convegno internazionale di storie della logica*, 19-30.
- Borg, Jana, Walter Sinnott-Armstrong, Vince Calhoun, and Kent Kiehl. 2011. “Neural Basis of Moral Verdict and Moral Deliberation.” *Social Neuroscience* 6 (4): 398–413. <https://doi.org/10.1080/17470919.2011.559363>
- Brzović, Zdenka. 2018. “Natural Kinds.” In *Internet Encyclopedia of Philosophy*. Accessed May 2, 2023. <https://www.iep.utm.edu/nat-kind/>

¹¹ According to Smith, this term was coined by J.E.J. Altham. For a discussion of the term, see Smith (1994, 118-125).

- Campbell, Richmond. 2007. "What is Moral Judgment?" *The Journal of Philosophy* 104 (7): 321-49.
- Carroll, Noël. 2012. "Recent Approaches to Aesthetic Experience." *Journal of Aesthetics and Art Criticism* 70 (2): 165-77.
- Carroll, Noël. 2015. "Defending the Content Approach to Aesthetic Experience." *Metaphilosophy* 46 (2): 171-88.
<https://doi.org/10.1111/meta.12131>
- Cushman, Fiery and Liane Young. 2011. "Patterns of Moral Judgment Derive from Nonmoral Psychological Representations." *Cognitive Science* 35 (6): 1052-75. <https://doi: 10.1111/j.1551-6709.2010.01167.x>
- Dancy, Jonathan. 2018. Moral Particularism. In *The Stanford Encyclopedia of Philosophy*, edited by Edward N. Zalta. Accessed June 15, 2022.
<https://plato.stanford.edu/archives/win2017/entries/moral-particularism/>.
- Decety, Jean and Thalia Wheatley, eds. 2015. *The Moral Brain: A Multidisciplinary Perspective*. Cambridge, MA: MIT Press.
- Decety, Jean, and Jason Cowell. 2014. "Friends or Foes: Is Empathy Necessary for Moral Behavior?" *Perspectives on Psychological Science* 9 (4): 525-37. <https://doi/10.1177/1745691614545130>
- Dreier, James. 1996. "Accepting Agent Centred Norms: A Problem for Non-Cognitivists and a Suggestion for Solving it." *Australasian Journal of Philosophy* 74 (3): 409–22.
- Fiske, Alan, and Taze Rai. 2014. *Virtuous Violence: Hurting and Killing to Create, Sustain, End, and Honor*. Cambridge: Cambridge University Press.
- Flanagan, Owen. 1993. *Varieties of Moral Personality*. Cambridge, MA: Harvard University Press.
- Foot, Philippa. 1967. "Abortion and the Doctrine of Double Effect." *Oxford Review* 5: 5-15.
- Gibbard, Allan. 2003. *Thinking How to Live*. Cambridge, MA: Harvard University Press.
- Gilligan, Carol. 2016. *In a Different Voice: Psychological Theory and Women's Development*. Cambridge, MA: Harvard University Press.
- Glasgow, Joshua. 2013. "Does Direct Moral Judgment Have a Phenomenal Essence?" *Journal of Moral Philosophy* 10 (1): 52-69.
<https://doi.org/10.1163/174552412X628841>
- Goodwin, Geoffrey, and John Darley. 2008. "The Psychology of Metaethics: Exploring Objectivism." *Cognition* 106 (3): 1339-66.
<https://doi.org/10.1016/j.cognition.2007.06.007>
- Greene, Joshua, R. Brian Sommerville, Leigh Nystrom, John Darley, and Jonathan Cohen. 2001. "An fMRI Investigation of Emotional

- Engagement in Moral Judgment.” *Science* 293 (5537): 2105-08.
<https://doi.org/10.1126/science.1062872>
- Greene, Joshua, Leigh Nystrom, Andrew Engell, John Darley, and Jonathan Cohen. 2004. “The Neural Bases of Cognitive Conflict and Control in Moral Judgment.” *Neuron* 44 (2): 389-400.
<https://doi.org/10.1016/j.neuron.2004.09.027>
- Greene, Joshua. 2008. “*The Secret Joke of Kant’s Soul*.” In *Moral psychology, Vol. 3. The Neuroscience of Morality: Emotion, Brain Disorders, and Development*, edited by Walter Sinnott Armstrong, 35–80. Cambridge, MA: MIT Press.
- Greene, Joshua. 2013. *Moral Tribes: Emotion, Reason, and the Gap Between Us and Them*: Penguin Press.
- Greene, Joshua. 2015a. “The Cognitive Neuroscience of Moral Judgment and Decision-making.” In *The Moral Brain: A Multidisciplinary Approach*, edited by Jean Decety and Thalia Wheatley, 197-220. Cambridge, MA: MIT Press.
- Greene, Joshua. 2015b. “The Rise of Moral Cognition.” *Cognition* 135: 39-42. <https://doi.org/10.1016/j.cognition.2014.11.018>
- Haidt, Jonathan. 2012. *The Righteous Mind: Why Good People are Divided by Politics and Religion*. New York: Pantheon Books.
- Hare, Richard. 1981. *Moral Thinking: Its Levels, Methods, and Point*. New York: Oxford University Press.
- Kant, Immanuel. 1998. *Groundwork of the Metaphysics of Morals*. Translated and edited by Mary Gregor. Cambridge: Cambridge University Press.
- Kelly, Daniel, Stephen Stich, Kevin Haley, Serena Eng and Daniel Fessler. 2007. “Harm, Affect, and the Moral/Conventional Distinction.” *Mind and Language* 22 (2): 117-31.
<https://doi.org/10.1111/j.1468-0017.2007.00302.x>
- Kumar, Victor. 2015. “Moral Judgment as a Natural Kind.” *Philosophical Studies* 172 (11): 2887-2910. <https://doi.org/10.1007/s11098-015-0448-7>
- Kumar, Victor. 2016a. “Psychopathy and Internalism.” *Canadian Journal of Philosophy* 46 (3): 318-45.
<https://doi.org/10.1080/00455091.2016.1165571>
- Kumar, Victor. 2016b. “The Empirical Identity of Moral Judgment.” *The Philosophical Quarterly* 66 (265): 783-804.
<https://doi.org/10.1093/pq/pqw019>
- Marek, Scott, Brenden Tervo-Clemmens, Finnegan Calabro, et al. 2022. “Reproducible Brain-wide Association Studies Require Thousands of Individuals.” *Nature* 603 (7902): 654-60.
<https://doi.org/10.1038/s41586-022-04492-9>

- Margoni, Francesco, and Luca Surian. 2020. "Question Framing Effects and the Processing of the Moral-conventional Distinction." *Philosophical Psychology* 34 (1): 76-101.
<https://doi.org/10.1080/09515089.2020.1845311>
- Martin, Wayne. 2006. *Theories of Judgment: Psychology, Logic, Phenomenology*. Cambridge: Cambridge University Press.
- May, Joshua and Victor Kumar. 2018. Moral Reasoning and Emotion. In *The Routledge Handbook of Moral Epistemology*, edited by Karen Jones, Mark Timmons, and Aaron Zimmerman, 139-56. London: Routledge.
- Moore, George Edward. 1903. *Principia Ethica*. Cambridge: Cambridge University Press.
- Nucci, Larry, and Elliot Turiel. 1978. "Social Interactions and the Development of Social Concepts in Preschool Age Children." *Child Development* 49 (2): 400-07.
<https://doi.org/10.2307/1128704>
- Osterloff, Emily. Nd. Brontosaurus: Reinstating a prehistoric icon. *Natural History Museum*. Accessed January 24, 2023.
<https://www.nhm.ac.uk/discover/brontosaurus-reinstating-a-prehistoric-icon.html#:~:text=That%20was%20until%20a%20study,status%20of%20this%20iconic%20dinosaur>
- Prinz, Jesse. 2007. *The Emotional Construction of Morals*. New York: Oxford University Press.
- Rehren, Paul and Walter Sinnott-Armstrong. 2022. "How Stable are Moral Judgments?" *Review of Philosophy and Psychology*,
<https://doi.org/10.1007/s13164-022-00649-7>
- Sackris, David. 2022. "The Disunity of Moral Judgment: Implications for the Study of Psychopathy." *Philosophical Psychology*.
<https://doi.org/10.1080/09515089.2022.2155125>
- Sackris, David, and Rasmus Rosenberg Larsen. 2020. "A Consideration of Carroll's Content Theory." *Journal of Value Inquiry*, 54: 245-55.
<https://doi.org/10.1007/s10790-019-09693-6>
- Sackris, David, and Rasmus Rosenberg Larsen. 2022. "The Disunity of Moral Judgment: Evidence and Implications." *Philosophical Psychology*. <https://doi.org/10.1080/09515089.2022.2056437>
- Sackris, David, and Rasmus Rosenberg Larsen. (2023). "Are There 'Aesthetic' Judgments?" *Erkenntnis*.
<https://doi.org/10.1007/s10670-023-00663-w>
- Sayre-McCord, Geoff. 2014. Metaethics. In *The Stanford Encyclopedia of Philosophy*, edited by Edward N. Zalta. Accessed June 15, 2022.
<https://plato.stanford.edu/archives/sum2014/entries/metaethics/>.
- Shafer-Landau, Russ. 2015. *The Fundamentals of Ethics*. Oxford: Oxford University Press.

- Seppälä, Selja, Alan Ruttenberg, and Barry Smith. 2017. “Guidelines for Writing Definitions in Ontologies. *Ciência Da Informação* 46 (1): 73-88.
- Sinnott-Armstrong, Walter, and Thalia Wheatley. 2012. “The Disunity of Moral Judgment and Why it Matters in Philosophy.” *The Monist* 95 (3): 355-77. <https://doi.org/10.5840/monist201295319>
- Sinnott-Armstrong, Walter, and Thalia Wheatley. 2014. “Are Moral Judgments Unified?” *Philosophical Psychology* 27 (4) 451-74. <https://doi.org/10.1080/09515089.2012.736075>
- Smith, Michael. 1994. *The Moral Problem*. Oxford: Blackwell Publishing.
- Stich, Steven. 2006. “Is morality an elegant machine or a kludge?” *Journal of Cognition and Culture* 6 (1-2): 181-89. <https://doi.org/10.1163/156853706776931349>
- Thomson, Judith. 1976. “Killing, Letting Die, and the Trolley Problem.” *The Monist* 59 (2): 204-17. <https://doi.org/10.5840/monist197659224>
- Verplaetse, Jan. 2009. *Localizing the Moral Sense: Neuroscience and the Search for the Cerebral Seat of Morality., 1800-1930*. New York: Springer. <https://doi:10.1007/978-1-4020-6322-0>
- Young, Liane, and James Dungan. 2012. “Where in the Brain is Morality? Everywhere and Maybe Nowhere.” *Social Neuroscience* 7 (1): 1-10. <https://doi:10.1080/17470919.2011.569146>
- Witherell, Carol, and Carolyn Edwards. 1991. “Moral Versus Social-conventional Reasoning: A Narrative and Cultural Critique.” *Journal of Moral Education* 20 (3): 293-304. <https://doi:10.1080/0305724910200306>

