**Chapter One**

This first chapter of the dissertation gives an original account of the normativity of concepts and compares it to other accounts of the normativity of concepts that have been introduced in the literature. The section will proceed by introducing and defending a definition, given in terms of necessary and sufficient conditions, for some concept's counting as normative.

**I.A Distinction between Normative Concepts and Properties**

Before introducing the definition of a normative concept, it is important to clarify a distinction between normative concepts and so-called normative properties, where these are both terms of art in my system. Normative concepts are invoked by metaethical theorists of all stripes to characterize the referents of ethical terms and other terms that are often marked out as normative, for example, "beauty." Both realists and anti-realists use normative terms like "good," "right," "ought," "should," "reason," and "value," and are thus committed to the use of the corresponding normative concepts. It is fair to call these terms and the corresponding concepts normative due to a distinction between descriptive language and normative language that I am claiming does not carve any substantive, corresponding distinction between the referents of descriptive language, according to some accounts, and the referents of normative language.

Descriptive language is language that describes the way the world is. Normative language is language that at least purports to partly describe the way the world should be, which can be spelled out in terms of its conformity to a possibly nonexistent norm or standard. It may be objected to the statement that normative language describes how the world should be that this claim uses a normative term to explain what a normative term is. Notably, it is controversial to hold that it is possible to give a completely reductive analysis of all normative terms in descriptive language, although all normative language apparently has an evaluative aspect that distinguishes it from purely descriptive language.

On some accounts, normative terms may still refer to properties that could also be picked out by purely descriptive language. For example, on a plausible anti-realist account of moral terms, "right" could refer to properties of the feeling of pleasure you experience when you do a good deed, such as giving money to the homeless. Combined with one plausible account from the philosophy of mind of what feelings of pleasure consist in, properties of the feeling of pleasure you experience when you do a good deed may just be identical to properties of the brain state you possess while you are performing the good deed. There is nothing about those properties themselves that makes them indescribable by a non-normative, descriptive phrase, such as "the relevant properties of that brain state that I had at time t," where time t is the time when I performed a good deed.

It may appear that the distinction between thin and thick moral terms, introduced by Williams, would be useful to introduce in connection with the distinction between descriptive and normative language. This is because thin moral terms and the corresponding concepts are thought to be paradigmatically normative. Because the distinction between thin and thick concepts concerns the natures of their referents and is only relevant to realist accounts of normative properties, however, this appearance is illusory.

A thin moral term, of which "good" for non-naturalist realists is an example, is defined as a moral term, no part of whose referent is descriptive. In contrast, a thick term like the noun "abuse" is

thought to refer to an item with both descriptive and normative aspects (Thanks to Erin Rodgers for the "abuse" example). That is to say, the referent of "abuse" in a given instance of the term's application is comprised of both descriptive and normative properties. The relevant normative property, in this example, is plausibly wrongness, considered as a normative property. It is impossible, according to conventional accounts of the meaning of "abuse," to call an action an instance of "abuse" without attributing wrongness to it. We see, however, that the referent of "abuse" also has a descriptive component. In referring to an action as an instance of "abuse," some description making reference to certain natural properties of the action is apparently used to fix that action as the reference of the term.

To see the irrelevance of describing terms like "good," according to moral antirealists, as 'thin,' consider that many anti-realists, unlike realists, hold that the term "good" has both a descriptive aspect and a normative one. Anti-realists are also committed to there being no intrinsically normative properties. Thus, even the normative aspect of a thin moral term, according to an anti-realist account of these, is exhausted by the evaluative function of the term as it is used in moral discourse.

The way I and those involved in the discussion of the distinction between thin and thick moral terms use the phrase "normative properties," these are not just identical to the properties picked out by normative concepts. Rather, normative properties, for my purposes, belong to a unique property class that is different in kind from the class of ordinary descriptive properties. To say that the class is "unique" is to say that no property belonging to the set of descriptive properties also belongs to the set of normative properties and that no property belonging to the set of normative properties also belongs to the set of descriptive properties. As Mackie says, "[i]f there were objective values, then they would be entities or qualities or relations of a very strange sort, utterly different from anything in the universe." (Mackie 1977 p. 38)

Now, go back to the earlier example on which "right" refers to "the relevant properties of that brain state that I had at time t." Even though properties fitting that description may be picked out by the normative concept "rightness," according to a certain metaethical theory, these properties still do not count as normative properties based on the distinction between descriptive and normative properties that I am drawing here. Normative properties are the properties that Mackie calls metaphysically queer, and do not count as descriptive properties under any possible description because they belong to a unique class of nonnatural properties. They are properties of so-called objective values. The issue of the relation between the descriptive properties of an item to which a normative property is attributed and the attributed normative property is discussed further in the following passage from Mackie:

> "Another way of bringing out this queerness is to ask, about anything
> that is supposed to have some objective moral quality, how this is linked
> with its natural features. What is the connection between the natural
> fact that an action is a piece of deliberate cruelty – say, causing pain just
> for fun – and the moral fact that it is wrong? It cannot be an entailment,
> a logical or semantic necessity. Yet it is not merely that the two features
> occur together. The wrongness must somehow be 'consequential' or
> 'supervenient.'" (Mackie 1977 p. 41)

The quoted passage drives home the point that objective value properties are different in kind from any of the other descriptive properties there are. Mackie notes that a property like "wrongness," if

it existed, would seem to always co-occur with certain natural features, though it would be considered a distinct property in its own right. Mackie himself rejects moral realism at an ontological level, and ultimately denies that there are any normative properties in the sense I mean. Many realists, including Moore and most intuitionists, however, have adopted accounts on which there are metaphysically queer properties or objective value properties. These are the "normative properties" under discussion here. Normative properties, on the accounts that posit these, are always picked out by normative concepts. Moreover, because normative properties in the relevant sense comprise a unique class, it is impossible to refer to them using descriptive concepts alone.

An objection to this point is that it is possible to be a moral realist and, at the same time, a reductive naturalist. On a reductive naturalist realist view, a normative concept is defined as being analytically equivalent to a descriptive concept. On some versions of the utilitarian view, for example, "good" is supposed to be analytically equivalent to "pleasure." Even if we grant that it is possible for normative properties to be picked out by descriptive concepts according to reductive naturalism, there is a question about whether so-called reductive naturalist realists challenge the dichotomy between descriptive and normative properties laid out here. The reply is that they do not, because they would not be positing the existence of normative properties in the sense specified. The universe within which they are working seems to consist of only descriptive properties.

To sum up the discussion so far, normative concepts are the concepts that are expressed by normative or evaluative terms. Since our concern is with the moral domain, it is with the normative concepts that correspond to moral terms. All metaethical theories that invoke the relevant terms, such as "right," "good," "ought," "should," "reason," and "value" invoke normative concepts. These concepts are contrasted with descriptive concepts, which are the concepts that are expressed by all terms that are not, for our purposes, normative terms. This distinction between descriptive and normative concepts does not map onto the distinction between thick and thin moral terms, which seems substantive exclusively for moral realists and, presupposing that there are normative properties, classifies moral terms along the dimension of the natures of their reference types (i.e. whether their referents are wholly or partially normative).

The distinction between descriptive and normative concepts also does not map onto the distinction between concepts used to refer to descriptive properties and those used to refer to normative properties. Descriptive properties are the properties of all the things we could encounter in the non-moral realm, whereas normative properties are the metaphysically queer properties of objective values. Normative concepts, for antirealists and reductive naturalist realists, can be used to refer to descriptive properties. Indeed, for antirealists and reductive naturalist realists about moral value, descriptive properties are the only properties there are in the universe for normative concepts to refer to. Antirealists and reductive naturalist realists also hold that descriptive concepts can be used to refer to those same descriptive properties picked out by normative concepts. This is the key difference between moral realists of the kind who posit objectively normative properties and antirealists and reductive naturalist realists as regards their use of normative concepts. Whereas antirealists and reductive naturalist realists hold that there is always some descriptive concept that is co-referential with any given normative concept, realists whose ontologies admit of normative properties deny that the referent of a normative concept can be alternatively picked out by any descriptive concepts.

**II.A Theory of Normative Concepts**

Notice that in characterizing normative terms and concepts, I included terms like "good," "right," "ought," "should," "reason," and "value" as examples. These terms in fact have been identified as normative by a wide range of writers on all sides of the metaethical spectrum. There remains, however, a question of why these terms in particular should be demarcated as belonging to a category distinct from that of the rest of language, which is descriptive. I hinted earlier that the answer probably has something to do with their evaluative role or function, but did not go into detail. Others have identified the prescriptivity of characteristically normative terms as a distinctive feature shared by many of these. This second part of the chapter defends necessary and sufficient conditions for some term's counting as normative that I believe are comprehensive and apply to all the terms and concepts, including those listed, that are generally thought to belong to this category.

The account being defended proposes two necessary and jointly sufficient conditions for some term or concept's classification as normative.

A term or concept is normative iff it:
1) Refers or attempts to refer to something thought to be capable of directing an agent's voluntary behavior in an instrumental way.
2) Is used to attribute value or a relation to some entity said to have value to its supposed referent.

To make some preliminary clarifications, these conditions are supposed to be necessary and jointly sufficient for classifying a concept as normative according to any metaethical view. They do not purport to give a reductive analysis of ontological normativity in terms of purely descriptive language. They also do not purport to give a reductive analysis of normative concepts in terms of purely descriptive concepts, since 2) appeals to value properties and the referents of value terms in the analysis. As mentioned previously, many have called into question the possibility of giving a completely reductive analysis of the normative domain, in language or metaphysics, in completely descriptive terms. On nonnaturalist realist views, this is generally assumed to be impossible due to the supposed existence of nonnatural properties, which are objective value properties. On naturalist realist views and antirealist views, everything 'said to have value' is ultimately comprised of descriptive properties, which suggests the possibility of an analysis of the normative in descriptive language alone.

Although 2), again, makes reference to value properties and the referents of value terms, I do not presuppose a value-first view of normative entities in including 2) in the analysis and view the analysis as being neutral with respect to so-called 'X-first' debates. X-first debates, to use the terminology of Erroll Lord, concern the issue of which normative entities are explanatorily basic (Lord 2023 p. 250). Even if, as reasons-first theorists maintain, the notion of a normative reason is explanatorily basic in analyses of from where all other normative notions derive their status as normative, it may still be the case that according to any metaethical view, a normative concept is used to attribute value or a relation to some entity said to have value to its supposed referent, and, that when a concept meets this condition and also refers or attempts to refer to something thought to be capable of directing an agent's voluntary behavior in an instrumental way, its satisfaction of these two conditions is sufficient for its classification as a normative concept. This analysis does not purport to

locate the most basic or fundamental normative concept or entity in any way. It provides a basis for categorizing any given concept as normative that is inclusive of all metaethical views so that the dissertation can later make the bold claim that W.D. Ross's concept of rightness would not be considered a normative concept, regardless of one's view about what the nature of normativity consists in.

There have been many other discussions of normativity in philosophical writing, and these two conditions may seem to call for further defense in light of some of them. To establish to the best of my ability that these two conditions are necessary and sufficient conditions for some concept's normativity, I first clarify the conditions briefly and run through my rationales for using specific words to frame them. Second, I show that the aforementioned, listed, paradigmatically normative terms, such as "good," "right," "ought," "should," "reason," and "value" all meet these conditions. Finally, I address proposed counterexamples; first to the claim, of each condition, that it is not necessary for the normativity of a concept, and second to the claim that the two conditions, jointly, are not sufficient for the normativity of a concept.

**II.A. The First Condition**

The first condition says that it is necessary that a normative concept refer or attempt to refer to something thought to be capable of directing an agent's voluntary behavior in an instrumental way. There are four implicit commitments about the referent of a normative term involved in the statement of this condition that need to be addressed: (A) that attempted, and not necessarily successful reference is enough to satisfy it, (B) that it states that the referent of a normative concept needs only to be thought capable, rather than actually capable of directing an agent's behavior, (C) that the relevant behavior must be voluntary, and (D) that the behavior that the referent is supposed capable of directing must be directed in a way that is "instrumental."

To begin with (A), consider the point that attempted reference is enough to satisfy (1) as a necessary condition for some concept's normativity. That is, successful reference is not implied by the language of the condition. The "attempted" is there so as not to exclude moral error theorists, who hold that moral terms all fail to refer due the non-existence or fictionality of their attempted referents. Error theorists still invoke morally normative concepts, and so fictional entities or non-entities, reference to which is merely attempted but does not succeed, should be included among their possible referents as these are outlined by condition (1). Matti Eklund is one person who has argued that it is possible for a community to use normative vocabulary in discourse that does not ascribe normative properties, either because its terms pick out properties that are normatively insignificant or because there do not exist normatively significant properties at all (Eklund 2017 14-15). Christine Korsgaard is another theorist who makes the point that error theorists, as well as everyone else involved in theorizing the normative domain of discourse, make use of normative concepts, even when they argue that those concepts fail to refer.

> "Normative concepts like right, good, obligation, reason, are our names
> for the solutions to normative problems, for what it is we are looking for
> when we face them. And if we sometimes succeed in solving those

6

problems, then there will be normative truths: that is, statements which employ normative concepts correctly. So it is true that the assumption of a realm of inherently normative entities or objective values is not needed to explain the existence of normative concepts, or the resulting existence of a category of normative truths. It is not because we notice normative entities in the course of our experience, but because we are normative animals who can question our experience that normative concepts exist." (Korsgaard 1996, p. 47)

Next, (B) says that the language of (1) implies that the referent of a normative concept does not need to be capable of actually directing an agent's behavior so long as it is thought to be capable of doing so. One rationale for including this caveat is the same as that for (A); if there are no reasons at all and the concept "reason" fails to refer to anything, then the concept does not refer to something actually capable of directing anyone's behavior. We might still want, however, to keep "reason" on the list of normative concepts according to this kind of error theory, and in a sense, (B) allows this. It seems like whatever fictional entity is picked out by "reason," according to such a theory, is *thought* capable of directing an agent's behavior, even if it is not *actually* capable of doing so.

Another reason to include the "thought" before capable in (1) is to accommodate cases that have been called cases of "true irrationality" (Korsgaard 1986 p. 12). Some of these cases are ones in which an agent has a reason to do something (e.g. Smith having a reason to tighten a particular screw in order to fix her bike), in the sense that she has a desire whose satisfaction would be causally contributed to by taking a particular action (e.g. a desire to bike to campus), but is motivated by her reason to do something different than the thing she has a reason to do (e.g. tightening a screw other than the one that required tightening, perhaps due to a mechanical error or lapse of memory). A similar point concerns those who argue that reasons may still be possessed by people suffering from moral accidie. In a case involving an agent in possession of a reason who suffers from depression, for example, the relevant reason might be *thought* capable of motivating her behavior without actually being capable of doing so due to the agent's condition (Dancy 1993 p. 5).

There may be a further question about why a normative notion must refer or attempt to refer to an entity capable of motivating an agent's behavior in the first place, rather than one that just "motivates" or "always motivates." Many other writers on both epistemic and moral normativity (including Adler 2002 (p. 51), Korsgaard 1986 (p. 11), Rinard 2019 (p. 3)) have defined normative reasons at large as they are defined here: as considerations that count in favor of acts or attitudes and that are capable of motivating agents to perform those actions or form those attitudes. A consideration that counts in favor of an action or attitude may be thought of as providing some justification for it, where some have tried to explain when this justification obtains in terms of a relation of fittingness between a normative reason and the action or attitude to which it is said to lend justificatory support (Lord 2023 p. 253).

The term 'capable,' as it is used to state what a normative reason is, most often seems to concern causal possibility. If it is impossible for a consideration, by way of an appropriate causal process, to lead to motivation for an action it is alleged to support, then it does not properly count as a

normative reason for acting in that way. The requirement that normative reasons be "capable" of motivating seems to be added for two main reasons.

On one level of explanation, it makes sense to say that normative reasons are capable of motivating agents but do not necessarily do so or do so simpliciter because it is clear from empirical evidence that normative reasons are not thought to motivate agents in every case where they are possessed. Even motivational internalists, who hold that an agent's possession of a reason guarantees her motivation in accordance with it in some fairly strong sense of "guarantees," must explain empirical data about agents who apparently possess reasons but lack motivation. This data is notable in cases belonging to the two kinds mentioned earlier: those of agents suffering from moral accidie and those of agents who act in cases of true irrationality. Korsgaard 1986 (p. 15), who has also noticed that cases like these might be thought by opponents to raise a challenge to motivational internalism as a theory, seeks to explain how internalists can accommodate the data they provide by clarifying what it is that motivational internalists actually require. She says that internalists do not require that rational considerations always succeed in motivating us, but only that they motivate us insofar as we are rational (Korsgaard 1986 (p. 15)). That is one sense in which supposed normative entities need only be thought *capable* of motivating.

On a different level of explanation, the "capable of motivating" language is called for by what are thought to be limits on the scope of voluntary behavior. If what counts as normative is related to the question of what agents ought to be motivated by and what an agent ought to be motivated by is constrained by what she is capable of being motivated by, then it makes sense to stipulate that the referents of normative concepts are thought capable of motivating behavior rather than incapable of doing so. This point may seem superfluous because it does not seem like many would hold that the referents of normative concepts should be thought incapable of motivating behavior. It is more revealing, however, when a qualifier about the voluntary status of the behavior is added. If what one ought to be motivated by in a voluntary way is constrained by what one *can* be motivated by in a voluntary way, then saying that a normative concept must refer or attempt to refer to something thought to be capable of directing an agent's *voluntary* behavior places a constraint on the kinds of things that can belong to the extension of a normative concept. What is incapable of motivating an agent's behavior in a voluntary way is a natural law or a coercive force (regardless of origin), and Adler 2002 (p. 51) makes a similar point as part of a text on epistemic reasons.

This brings us to (C): that the behavior that the referent or supposed referent of a normative concept is thought to be capable of motivating must be voluntary. Most theorists about moral normativity have stipulated that behavior motivated by a normative reason is "voluntary" in order to distinguish whatever it is normative concepts attempt to refer to from the forces of coercion. Without the addition of "voluntary" to 1), moreover, concepts of many bodily processes and the mechanisms underlying those would also count as normative, as would the concepts of certain coercive forces and coerced actions insofar as these were said to have value.

H.L.A. Hart stresses a similar distinction between the concept of law and the concept of coercive force in responding to a theory according to which laws are supposed to amount to coercive orders (Hart 1961 p. 82). His objection to the theory is that it does not capture well the notion of legal obligation, as we would not say of a man being ordered at gunpoint to hand over his money that he had an obligation to do so. If we think of laws as supporting legal obligations, then, the concept of law must

attempt reference to something whose authority over behavior is based on more than the features of a coercive situation. Hart and others have also noted the importance of distinguishing the motivational justification for an action (Copp 1995 p. 187, Falk 1947 p. 122, Hart 1961 p. 84, Korsgaard 1996 p. 13, 52, 60), which may consist in the effect of a coercive force in inducing an agent to perform it, from its normative justification. The normative justification for an action essentially has to do with the agent's justification, from a first-personal perspective, for voluntarily engaging in the relevant activity (Hart 1961 p. 90, Korsgaard 1996 p. 16).

There is a question about whether the requirement that the referent of a normative concept be capable of directing behavior in a *voluntary* way resembles another suggestion that others have made about normative behavior, namely that in order to count as normative, behavior must be performed on the basis of reasons that have second-order reflective endorsement from the agent for whom they are reasons (Korsgaard 1996 p. 17). The voluntariness requirement is similar if actions performed voluntarily are always performed on the basis of reasons that have second-order reflective endorsement, although it is not clear that there is a logical equivalence between voluntary behavior and behavior performed from reasons that have second-order endorsement. For now, it is enough to say that these two requirements are not identical.

Finally, there may be a question about whether motivated behavior is ipso facto voluntary, and if it is not, whether the addition of 'voluntary' to 1) makes the condition controversial. Voluntarist views in Korsgaard's terminology, for instance, may be thought to challenge the requirement that normative concepts always "refer or attempt to refer to something thought to be capable of directing an agent's voluntary behavior" if reasons for action borne out of coercive situations are thought, by definition, to be incapable of leading to behavior that is truly voluntary. This is because voluntarism holds that obligations arise from the command of an authority, who is sometimes thought to be a religious deity (Korsgaard 1996 p. 18). The more general question is about specifying the sense of 'voluntary' used in the statement of (1).

There may be a background assumption that truly voluntary behavior is such that engaging in it would satisfy or be a means to satisfying a desire possessed by an agent independently of any external influences on her final ends. When an agent is in a coercive situation, in other words, her cognition of the coercive dynamic of the situation is a consideration that gives her a new desire that she didn't have before: a desire to avoid whatever consequences would arise from her violation of the sanction or her resistance to the forces of coercion. Satisfying the negative desire that arises from the situation of an agent who is subject to a sanction or coercion is arguably not cognized by her in deliberation as a means to the satisfaction of any other desires she has in a straightforward way. Rather, it seems like avoiding the consequences of trespassing a sanction or defying a coercive force is the object of a new intrinsic desire of hers that she possesses only in virtue of the situation she is in. Coerced behavior may also be involuntary because coercive situations can give rise to impulsive action, or unreflective action undertaken without consideration of whether it coheres with one's value system (Davis 1984 p. 189). Arguably, impulsive action is also involuntary in virtue of being unreflective.

That coercive situations can introduce new intrinsic desires for agents or result in their performance of impulsive actions undergirds the commonly made assumption that reasons given by threats of punishment and coercive forces alone do not count as normative. Namely, these features of coercive situations reveal the sense in which involuntary, coerced actions are performed against the will

of the agent who acts. If the uncontroversial intuition that reasons for actions performed 'against one's will' are not really normative is correct, then it is appropriate to require that normative concepts refer or attempt to refer to things thought to be capable of directing only an agent's 'voluntary' behavior. Derek Parfit has also described the behavior that is responsive to reasons for action as 'voluntary' (Parfit 2011 p. 48).

The final aspect of (1) that calls for explanation is (D), that the behavior that the referent or supposed referent of a normative concept is thought capable of directing be directed in a way that is "instrumental." This adds that the behavior directed by the referent or supposed referent of a normative concept is end-directed, or in some way goal oriented. Even where no substantive state of affairs is marked out as the end of behavior that is normatively justified, many theorists have held that all practical reasons constitutively involve a desire state (Davidson 1963, Hursthouse 1991). Because the desire state that partly constitutes an agent's practical reason is supposed to be satisfied when the agent acts on the reason, according to the relevant theories, the agent's behavior in acting on the practical reason is instrumental to the satisfaction of the relevant state. This remains true even when goal-directed behavior is intrinsically motivated, or wanted for its own sake. If all normative reasons are practical and acting on practical reasons always involves the satisfaction of a desire state, then the behavior that the referent or supposed referent of a normative concept is thought capable of directing is always instrumental in the specified sense.

There are two points that might be thought to challenge the claim that the concept of normative behavior essentially involves instrumentality. First, one could say regarding the concept of an intrinsically good activity that it is not clear that the behavior it attempts to refer to is directed in a way that is instrumental. The reply to this point is that the behavior that the concept of an intrinsically good activity attempts to refer to is instrumental to realizing a valuable state of affairs. Concepts such as those of right acts refer or attempt to refer to act types that are capable of motivating agents to instantiate those types in behavior voluntarily, and in such a way that performing the action is instrumental to instantiating the act type.

Second, if we think of emotions as themselves having the status of normative reasons in some cases, it is not clear that actions motivated by occurrent emotion can be said to "satisfy" the emotions that constitute their reasons in any instrumental way. Someone who retaliates angrily in response to a stranger's harassment in the street, for example, might find that she remains angry after she has been motivated by anger to act in a retaliatory way. This is in contrast to the way someone who is motivated to act by a desire is thought to have satisfied the desire once she has acted on it, in such a way that the agent no longer possesses the motivating desire state afterwards.

There are two ways to address this objection. The first is that the thesis that emotions in themselves constitute normative reasons for action has not been widely defended or maintained. The second is that if emotions do in themselves constitute normative reasons for action, this could only be true either because acting from them realizes a state of affairs with intrinsic value or because emotion states encode information about what would be instrumentally valuable to the agents who possess them. If occurrent emotions are normative reasons, then based on whatever justification one endorses for maintaining this, acting on them as normative reasons is instrumental to realizing value. Intrinsic value concepts themselves do refer or attempt to refer to something capable of directing behavior in an instrumental way.

**II.B. The Second Condition**

We may now move on to the second necessary condition for some concept's classification as normative, which, when satisfied with the first, is also sufficient. Recall that condition (2) says that a normative concept is used to attribute value or a relation to some entity said to have value to its supposed referent. Two points that call for linguistic clarification, here, are (E) why the condition does not simply stop at "used to attribute value" but includes "or a relation to some entity said to have value" and (F) why "supposed referent" is still included. General controversy surrounding whether (2) is really a necessary condition for attributing normativity to a concept will be addressed later with other counterexamples.

Regarding (E), this addition, on one hand, allows the account of normative concepts to remain neutral with respect to any metaethical position. I say that normative terms can attribute relations to items "said to have value" in lieu of saying they attribute value in order to clarify that I am not requiring the use of normative terms to actually attribute value properties of the kind posited by moral realists to their referents. The term "good" for expressivists, for example, does not attribute a value property to its referent, though it still plausibly attributes a relation between its referent and some descriptive features that are said to have value according to expressivists (e.g. desires). It is fair to assume that according to an antirealist metaethical theory, all normative concepts merely attribute relations between their referents and things said to have value rather than actually attributing value properties to their referents.

A second reason to hold that the attribution of a relation to an entity said to have value to the referent of a concept can be enough to make that concept "normative" is that even among realists, certain normative terms are thought to derive their status as such from the relations of their referents to items said to have value, despite not being thought themselves to attribute intrinsic value properties to their referents. To give a concrete example, some utilitarians do not hold that they are attributing intrinsic value to an action in describing it as "right." We might still want to say, however, that the relevant utilitarians would hold that "right" is a normative concept according to this kind of utilitarian view. One way of explaining why might be to point out that "right" attributes instrumental value to actions according to utilitarianism. Mill's view is exceptional in that he says people come to view right actions as parts of happiness, and thus as intrinsically valuable, in virtue of their association with the production of happiness as an end (Mill 1861 p. 212). On other versions of utilitarianism, however, instrumental value does not often seem conceived as a property possessed by actions apart from their relations to sources of intrinsic value. Attributing instrumental value to something just consists in attributing to it a relation to some other item said to have intrinsic value.

(F) can be addressed fairly briefly by commenting that "supposed referent" is still added to accommodate error theories on which networks of normative terms may still be thought to exist. Though some term in question on such a theory may fail to refer, its use may still involve attempted reference-fixing via the attribution of a relation between its supposed referent and that of another term. Although that other term may also have an empty referent, this would nonetheless count as something "said to have value" according to the relevant sort of theory. The former concept would thus satisfy 2) by attributing a relation to some entity said to have value to its supposed referent.

An objection concerning (2) that was addressed more briefly earlier is that it is problematic to invoke the concept of value as part of an analysis of the normativity of concepts if we concede that the concept of value is itself a normative concept. The objection, again, is that invoking the concept of value as part of the analysis makes the analysis nonreductive, or incapable of accounting for the normativity of a concept by using only descriptive concepts. Relatedly, if the concept of value is itself normative, it seems like no reductive analysis of its normativity can be given by the foregoing analysis, even one given entirely in terms of other normative concepts.

To introduce a more detailed reply to this objection, recall that the second condition, in simpler terms, stipulates that the referent of a normative concept bears some kind of relation to the referent of a value concept as a matter of conceptual necessity. On any given metaethical view, a basic value concept may either refer to an objective value property or a property considered descriptive.

If a basic value concept refers to an objective value property, the normativity of the concept of value is accounted for by the attribution of that particular property, and the normativity of other normative concepts is accounted for by their use to attribute relations to that particular property to their referents. It is true that for realists, the foregoing analysis is nonreductive, but arguably, realists must hold that no truly reductive analysis of the normativity of concepts, in terms of descriptive concepts alone, is possible. This is a feature of their view that there are objectively normative properties.

If a basic value concept refers to a descriptive property, on the other hand, the normativity of the concept of value is also accounted for by the attribution of that particular property. Antirealists might call something valuable in respect of its disposition to be chosen or some other relation it has to the attitude of valuing. The normativity of other normative concepts would then be accounted for by their use to attribute their relations to their referents of things said to have value, namely, those descriptive properties that a given antirealist account calls valuable. The analysis, as it applies to antirealist metaethical theories, is reductive in the sense that it allows us to explain why a concept is classified as normative by invoking only descriptive concepts.


## III. Challenges for the Theory

So far, the chapter has gone into detail about the language used to state the following necessary and jointly sufficient conditions for some concept's counting as normative:

A term or concept is normative iff it:
1) Refers or attempts to refer to something thought to be capable of directing an agent's voluntary behavior in an instrumental way.
2) Is used to attribute value or a relation to some entity said to have value to its supposed referent.

Some further explanation may be given to clarify that the paradigmatic normative terms, including "good," "right," "ought," and "reason" do satisfy these conditions. First, "good" on any metaethical theory does refer to attempt to refer to something thought capable of directing an agent's voluntary behavior in an instrumental way if we grant that people are sometimes thought to be motivated to behave for the sake of bringing about whatever the bearers of goodness are thought to be.

"Good" on realist theories is used to attribute value to its referent, and on antirealist theories is used to attribute a relation to some entity said to have value to its referent.

Second, "right" according to any metaethical theory refers or attempts to refer to something thought to be capable of directing an agent's voluntary behavior in an instrumental way if we grant that agents are sometimes thought to be motivated to act rightly by the conception of some behavior as right. Even if we do not grant this, the concept of rightness satisfies 1) if its referent is sometimes a basis for reactive behavior. "Right," on the standard accounts that occupy the conceptual landscape, is always used to attribute either goodness conceived as a property, a relation to goodness conceived as a property, or a relation to something said to be "good" to its referent.

Finally, "ought" and "reason" may be taken as a package deal, given that they are often treated as interdefinable. Michael Pendlebury, for example takes "x ought to A" to be logically equivalent to "There are good reasons for x to A" and takes "x must A" to be equivalent to "There are compelling reasons for x to A" (Pendlebury 2002 p. 188). Mackie endorses a version of the proposal that "a ought to G," for some agent a and action G, is equivalent to "there is a reason for a's G-ing." (Mackie 1977 p. 74). Hare holds that "ought" statements correspond to reasons by taking them to make different types of prescriptions, which are made in light of information about accepted principles and the facts in virtue of which they apply (Hare 1952 p. 155-6). Regarding the conditions on the normativity of a concept, reasons for action are uncontroversially thought to be capable of directing an agent's voluntary behavior in an instrumental way. On some normative theories, reasons themselves have intrinsic value, on others they have instrumental value, and on some others, they bear a relation to entities said to have intrinsic or instrumental value. If attributions of reasons are equivalent to "ought" statements, statements of obligation also satisfy the aforementioned conditions.

We have now reached the point where consideration of counterexamples to the theory is in order. First, I respond to the person who says 1) is not necessary because cases exist that show that normative entities need not be capable of directing behavior at all. One could also use counterexamples to push back on the point that the behavior a normative entity is thought capable of directing must be voluntary, and so it might make sense to say something about this. Next, I respond to the objection that 2) is not necessary because normativity is not essentially related to value.

If we assume reasons internalism, it looks like agents who exhibit "pure evil" (Joyce 1998 p. 21) or those resembling the figure Eklund terms the "Bad Guy" (Eklund 2017 p. 1) are counterexamples to 1). Reasons internalism is a view that was historically popularized by Williams and says that an agent has a reason to act in a particular way only if she has a desire whose satisfaction would be causally contributed to by acting in that way (Williams 1979 p. 101). As Eklund describes the "Bad Guy," this is "someone who does things and is motivated by bad desires...[,or someone] objectively out of sync with how to conduct one's life."(Eklund 2017 p.1) Joyce similarly describes the person who exhibits "pure evil" as one who has bad intentions and desires to act in ways that are contrary to accepted moral standards. (Joyce 1998 p. 21) For our purposes, here, we might think of the purely evil person as a sociopath. On standard accounts of desire's involvement in motivation, a desire to $\Phi$ (where $\Phi$ is an act type) is always what motivates an agent to $\Phi$ voluntarily. A sociopath is someone whose desire system is

intrinsically deficient in such a way that she is not capable of having desires, or consequently, desires that are capable of motivating her, to act in the ways that conventional morality recommends.

Because the figure who exhibits pure evil does not have the requisite desires to possess the reasons that most would consider normative, a reasons internalist would deny that the uncontroversially normative, moral considerations that count as reasons for others are reasons, or even considerations that count as normative, for her. This objection to 1) most directly challenges its claim that normative concepts always refer to things that are capable of directing an agent's behavior if we also assume a non-relativist form of moral realism (For Eklund's discussion of a similar case, see Eklund 2017 p. 3). The purely evil agent could be thought a counterexample to the claim that normative concepts always refer to entities capable of directing behavior for a non-relativist realist who is also a reasons internalist because there is a sense in which the moral considerations that count as normative moral reasons for everyone else are incapable of directing her behavior.

To see this, recall that we have supposed that the purely evil agent is a sociopath who is incapable of having desires to act in accordance with moral considerations corresponding to accepted moral standards for reasons having to do with her psychological constitution. If all voluntary behavior is motivated by the same desires that are presupposed by normative reasons, as some internal reasons theorists suppose, then the purely evil agent's inability to have desires to act in ways recommended by conventional moral considerations explains both why those considerations cannot count as reasons for her and why they are incapable of directing her voluntary behavior.

One thing to point out in responding to the idea that this case calls 1) into question is that moral realists are often not reasons internalists. It is, nonetheless, a logically possible position. One solution to the relevant case that preserves 1) is to concede relativism and hold that while the agent who exhibits pure evil would not consider normative the considerations that others take to be reasons, this is only the case because the standards that apply to everyone else do not apply to her. She has a separate normative vocabulary that is not coextensive with the regular person's normative vocabulary and that refers in accordance with the standards that are objectively binding on her, either owing to her domain or other facts that objectively make her subject to alternative standards.

Without abandoning the commitment against relativism, a different reply to the realist, reasons internalism case that preserves 1) as a necessary condition on a concept's classification as normative is one that has been endorsed by Korsgaard and Smith: that moral considerations that count as normative reasons for others are not normative for purely evil people because their constitution is such that they do not count as rational agents. For Korsgaard and Smith, moral considerations corresponding to accepted, and in some loose sense, 'objectively right,' moral standards are only capable of motivating agents and indeed, of being genuine reasons for them, insofar as those agents are rational (Joyce 1998 p. 22).

There is an interesting question that is beyond the scope of this section about whether the conception of rationality required to deflect the objection as Korsgaard and Smith would is substantive in a way that antirealists would find objectionable. It seems plausible that it is. Leaving that aside, it may be important to note of the purely evil agent case for realist, reasons internalists that neither the line of reply that holds moral reasons are capable of motivating only rational agents nor the line of reply that

assumes relativism would call into question the truth of the general claim that the concept of a reason is always normative. According to neither reply to the alleged counterexample does the purely evil agent have a "non-normative reason;" according to both, she simply has different reasons than everyone else.

Motivational judgment internalists may also see the way moral judgments motivate, according to their view, as calling into question that 1) is necessary for some concept's classification as normative. This is because motivational judgment internalists hold that agents cannot occurrently affirm a moral judgment, spelled out as a statement of obligation applying to a particular case, in thought without being motivated in accordance with its content to some degree. One way of explaining why affirming a judgment content guarantees corresponding motivation on a strong form of this kind of internalist view is that motivation is somehow built into the sense of a moral judgment qua linguistic item. As Korsgaard states the view, moral motives are "inevitably connected with the thought of duty." (Korsgaard 1996 p. 81)

Proponents of motivational judgment internalism might see their picture of the way moral judgments motivate as challenging 1) on the ground that not all moral concepts refer or attempt to refer to something thought to be capable of directing an agent's *voluntary* behavior. That is, the way in which a statement of obligation guarantees motivation is in some sense involuntary on the internalist picture just described. Insofar as agents affirm a moral judgment content, they have no choice in the matter of whether to be motivated in accordance with it. Behavior that issues from motivation guaranteed by thoughts with certain contents, about the possession of whose motives an agent has no choice, does not seem like hallmark 'voluntary' behavior, in other words. Yet internalists might still want to say that the concept of obligation or moral judgment is normative.

The reply to the claim that any case of moral motivation, according to motivational judgment internalists, provides a counterexample to 1) is to note that some judgment internalists, Korsgaard among them, have held that while certain motives necessarily accompany our affirmation of certain judgment contents, we have some choice about whether "we allow ourselves to be moved by the motives which morality provides." (Korgaard 1996 p. 81) Indeed, she says the question of whether we allow ourselves to be so moved is one way of stating what she calls "the normative question." Elsewhere, Korsgaard notes that the normative question is essentially about seeking justification over and above mere explanation for the motivational claims of morality on agents (Korsgaard 1996 p. 16-17). That explanation of motives that succeeds in the absence of adequate theoretical justification is supposed to problematize the normativity of a moral theory, for Korsgaard, is consistent with the demand that the behavior directed by the referent or supposed referent of a normative concept be voluntary.

A different way to respond to the argument that judgment internalism problematizes 1) is to hold that insofar as motivational internalism portrays behavior in accordance with moral judgments as involuntary, this is a data point that counts against adopting this picture of what is involved in being motivated in accordance with a moral judgment. There appears to be ample empirical evidence of cases in which agents who sincerely believe moral judgment contents are not actually moved in accordance with those contents. The view that corresponding moral motivation is some sense involuntary where moral belief exists is therefore untenable.

Opponents could raise certain non-moral, normative concepts as counterexamples to the claim that 2) is a necessary condition for the normativity of a concept. Concepts such as legal obligation and

15

the thin terms "good" and "right" as they are used in games to refer to objectives and possible or actual behavior might be introduced as normative concepts that do not attribute value their referents or supposed referents. To use a concrete case, someone objecting along these lines might say that when we claim that a person has a legal obligation not to jaywalk, we do not, by doing so, attribute any all-things-considered value to the referent of the term "obligation" as it is used in that instance. Whatever is referred to by "obligation" when we point out that individual's legal obligation not to jaywalk could have net disvalue, for example, if it induces him to obey the law and arrive late for an important meeting that would affect the welfare of many others.

I have already said that this account of normativity applies chiefly to moral concepts, but the account does also generalize to all practically normative concepts. To prevent the account from generalizing, we would only have to replace "value" in 2) with "moral value." The reply to the objection about other kinds of value concepts, and, namely, those that are associated with institutions such as legal systems and games, is that institution-specific normative concepts do attribute instrumental value to their referents. The referent of "his legal obligation not to jaywalk" does have value when the referent of that expression is itself considered as being instrumental to the maintenance of the legal system or the attainment of its institutional ends. This is not to say, however, that the legal obligation has intrinsic value or instrumental value relative to the personal ends of the agent in question. The same points apply to normative terms as they are used in contexts that are specific to particular games, in which players have ends distinctive to those games that are laid down by their rules.

Another point of contention may concern the claim that the concept of even moral obligation attributes value or a relation to something valuable to its supposed referent. It may be unclear why this would be required of the concept of obligation. To see why, note that something called an obligation is either a real normative entity, a fictional entity, or something descriptive that is referred to by that term. Regardless of to what the concept of obligation refers on a particular metaethical theory, an uncontroversial feature of the concept, according to any theory, is that its application by an agent can cause her to believe that a certain course of action is among the most choiceworthy in a particular situation, in such a way that would license her to take that course of action. If we believe actions to be worthy of choice only insofar as we believe that they are intrinsically or instrumentally valuable, then the point that applying the concept of obligation characteristically induces agents to beliefs about which actions would be choiceworthy implies that whatever "obligation" refers to is always related to something said (or equivalently, believed) to have value, namely, a choiceworthy action.

One could push back on this and say that the fact that the application of the concept of moral obligation in a given situation characteristically induces an agent to believe that a particular course of action is best does not imply that its referent is related to the value or purported value associated with the relevant action. This is a point about language. It could be something about the sense of the concept of obligation or its conceptual role when it is applied in a particular case, that is, that induces the characteristic belief rather than a relation between the referent of the concept and the action. This is a fair objection. One way to respond is to point out that when we communicate with each other using concepts, their referents are typically the things we are communicating about. Even in a case where a concept has a special role that is itself capable of inducing certain beliefs, the beliefs induced by that role are supposed to bear some relation to the actual referent of the concept.

To use a concrete example to illustrate this point, Eklund provides an account of slurs that gives an example of how a feature of language not related to its referent could induce in a speaker a belief about how to behave. This is an account of slurs which Eklund terms "The Tone View." The Tone View holds that the evaluative aspect of a slur is supposed to be accounted for by its conventional implicature (Eklund 2017 p. 67-68), which is a feature of the conventional use of a linguistic item rather than a feature of its sense or referent as these have been traditionally conceived. Eklund writes:

> "On one popular view, to use a slur for a given group of people is equivalent
> to using a non-normative word "--" ("black," "Jewish"...), but with a negative
> or contemptuous tone of voice; it is only that the negativity is conventionally
> associated with the word. Less metaphorically put, the evaluation can be held
> to reside not in what is *said,* in Grice's favored sense, but in some other feature
> of how what is said is communicated. Perhaps there is a conventional implicature:
> the speaker's use of the slur conventionally implicates that the speaker has a
> negative attitude toward –s. Conventional implicature is an aspect of meaning
> that does not directly affect reference." (Eklund 2017 p. 67-68)

Although conventional implicature does not affect reference, beliefs of a speaker about how to behave toward a particular person based on the use of a slur would imply a relation between the actual referent of the slur in that instance and the negative attitudes toward the relevant group conventionally implicated by its use. Similarly, even if the conceptual role or sense of the concept of obligation is what most directly explains why its application can give rise to an agent's beliefs about which course of action should be chosen, the beliefs themselves, insofar as they are based on the application of the concept of obligation, would imply a relation between the actual referent of the concept and whatever features of its sense or role allow its application to give rise to the relevant belief.

A final way to call into question the analysis of conceptual normativity given in this chapter would be to challenge that the two conditions involved are jointly sufficient for the normativity of a concept. This could involve finding a counterexample that shows there is a concept that meets both of the conditions, but still does not count as normative. The two conditions are restated here:

A term or concept is normative iff it:
1) Refers or attempts to refer to something thought to be capable of directing an agent's voluntary behavior in an instrumental way.
2) Is used to attribute value or a relation to some entity said to have value to its supposed referent.

The most obvious candidates for concepts that arguably meet both of these conditions but resist classification as normative concepts are emotion concepts and the concept of desire. Emotion concepts include the concepts associated with the terms "love," "hate" "anger," "joy," and more. These have already been discussed briefly in connection with the instrumentality language used in stating (1). A reply to the claim that emotion concepts are counterexamples to the joint sufficiency of (1) and (2) for conceptual normativity is to deny that emotion concepts satisfy (1), as some have held by claiming that the way in which emotions are capable of directing the voluntary behavior of agents is non-instrumental. That is, an agent can be directed by an occurrent emotion to behave in a voluntary way without her behavior in that instance being end-directed.

Another way to account for emotion concepts is to hold that they are normative. This is the line that some theorists seem forced to take with the concept of desire as well, although for others, there might be more of a basis for claiming that the concept of desire fails to satisfy (2), e.g. according to those who hold both that desires do not represent their objects as having value and that there is nothing intrinsically valuable about the satisfaction of a desire. The claim that emotion concepts and the concept of desire are normative concepts might be thought hard to defend because given the accounts of these concepts developed by those working in the philosophy of mind, their normativity could call into question whether there is really any substantive distinction between normative concepts and descriptive ones.

That emotion concepts and the concept of desire do not automatically strike us as 'definitely not normative,' however, is reassuring for the analysis. Additionally, if descriptive concepts are just those that describe natural phenomena, nothing has been said here to rule out the possibility of some concept's being both normative and descriptive. It does seem plausible that concepts such as those corresponding to emotion terms and the concept of desire occupy a gray area between normative and descriptive, and that different theories would classify them on different sides of the dividing line.

It is likely that proponents of theories that classify emotion concepts and the concept of desire as non-normative disagree with those who classify these concepts as normative in virtue of an underlying disagreement over whether these concepts really satisfy both (1) and (2). For some expressivists, holding that the concept of desire is normative is a natural option, since some expressivists propose that 'good' refers to the state of affairs that occurs when a desire of an agent's is satisfied. Thus, so long as someone with this view accepts an account of 'desire' on which it satisfies (1), (2) is satisfied in virtue of this other commitment of their theory.

Some non-naturalist realists, in contrast, might resist the classification of the concept of desire as normative by accepting a highly naturalistic account of the concept of desire that does not reveal how the referent of 'desire' could either have a non-natural property or be related to one of these. Non-naturalist realists hold that 'good' and the most basic value terms refer to non-natural properties. Because attributing value or a relation to something said to have value to an entity, for non-naturalists, involves attributing a non-natural property or a relation to something that possesses a non-natural property to it, it might be unclear to proponents of the non-naturalist view how a highly naturalistic concept of desire could ever satisfy (2). It might be clarifying to add to this that I am not claiming that non-naturalists are necessarily committed to denying that the concept of desire is normative. Assuming that the concept of desire always refers to desire states, a non-naturalist realist might concede that the concept of desire satisfies 2) either by rejecting the kind of highly naturalistic view of desire states that would disallow attributing non-natural properties to these or by granting that desire states can represent non-natural properties and thereby stand in relations to things said to have value.

Finally, it may be worth mentioning that the conditions of this account are only said to have been satisfied when they hold for every instance of a given concept's application, and not just a few instances. That the concept of belief can refer, in some instances, to something capable of directing an agent's behavior in a voluntary way and attribute to its referent a relation to something that has value

(i.e. when the content of the belief is a proposition that attributes value), does not make the concept of belief normative according to this account. For that to be the case, the concept of belief would have to satisfy (1) and (2) constitutively and as a matter of conceptual necessity; the application of a concept in some instance would have to satisfy both (1) and (2) to even count as an instance of the application of the concept of belief.

**Chapter Two**