

Is moral projectivism empirically tractable?

Richard Joyce

Penultimate draft of the article appearing in *Ethical Theory and Moral Practice* 12 (2009): 53-75.

“Projectivism” means different things to different philosophers. Even restricting our attention to *moral* projectivism merely identifies our subject matter while leaving the nature of the *projectivist* component of the thesis indeterminate. The objectives of this paper are to home in on and clarify one central thesis that seems deserving of the name “moral projectivism,” and to call attention to the fact that it is an empirical hypothesis and thus must be tested as such. I should at the outset immediately quell any expectations that in this paper I will design, develop, or even suggest any experimental methods. The preliminary task of identifying and clarifying a target hypothesis is sufficiently complicated to fill the paper, and I lack the space to propose any specific empirical procedures. If what follows serves to reorient thinking about moral projectivism in an empirical direction, if it encourages people to reflect on ways by which it might be properly tested, then I will be satisfied with the contribution.

1. The Many Moral Projectivisms

Along with just about everyone else who discusses the topic of projectivism, I shall begin with David Hume:

’Tis a common observation, that the mind has a great propensity to spread itself on external objects.
(1740/1978)

“A common observation”? Restricted to the domain of philosophers of the 17th and 18th centuries, yes. A century before Hume’s comment, Descartes had described humans as “accustomed...to attribute to bodies many things which belong only to the soul” (1641/1970: 109). A few years before that, Galileo declared that “many sensations which are deemed to be qualities residing in external subjects [including tastes, odors, smells, and heat] have no real existence except in ourselves, and outside of us are nothing but names” (1623/1960: 12). Hobbes very probably was a projectivist about many aspects of human experience, including morality (see Darwall 2000). And Newton (who was of course enormously influential upon Hume) endorsed what can be interpreted as a projectivist view of color in his best-selling *Opticks* of 1704.

In the above quote from the *Treatise*, Hume is discussing not morality but the human idea of *necessary connection*. It is generally assumed that he intends the same treatment for morals, beauty, color, sounds, and other perceptible sensory qualities. In the *Enquiry Concerning the Principles of Morals*, he apparently applies the “common observation” to morality:

Thus the distinct boundaries and offices of reason and of taste are easily ascertained. The former conveys the knowledge of truth and falsehood: the latter gives the sentiment of beauty and deformity, vice and virtue. The one discovers objects as they really stand in nature, without addition or diminution: the other has a productive faculty, and gilding or staining all natural objects with the colours, borrowed from internal sentiment, raises in a manner a new creation. (1751/1983)

However, R.M. Sainsbury (1998) takes issue with the assumption that what Hume says for necessary connection is supposed to go for morality and the rest. The difference, thinks Sainsbury, is that (according to Hume) our projected idea of necessary connection leads us to massive doxastic error about the nature of reality, whereas our projected ideas of vice and virtue (and moral qualities in general) do not necessarily result in error. Sainsbury, in fact, thinks that Hume's moral projectivism is compatible with a commitment to moral realism.¹ I am not here concerned with whether Sainsbury has correctly identified a distinction present in Hume's texts; I am interested in the general distinction between error-implying and non-error-implying versions of moral projectivism. It is a distinction that should be familiar to scholars of modern metaethics. On the one hand, John Mackie—a prominent advocate of the moral error theory (and who coined the label, no less)—is a moral projectivist (1977: 42-46; 1980: 72).² On the other hand, Simon Blackburn—a staunch critic of Mackie's error theory and promoter of the contrary metaethical theory of noncognitivism—is also a moral projectivist (1993; 1998).³

So we already have three positions in play: First, moral projectivism coupled with an error theory (Mackie); second, moral projectivism coupled with realism (Sainsbury's Hume); third, moral projectivism coupled with noncognitivism (Blackburn). A conspicuous question is whether these are really three distinct variants of moral projectivism, or whether we have a single projectivism that is neutral among these metaethical options.

Sainsbury evidently thinks that there is more than one kind of projectivism discernible in the Humean texts. He argues that the *Treatise* relation of “spreading” (pertaining to the idea of necessary connection) is distinct from the *Enquiry* relation of “gilding or staining” (pertaining to morals, inter alia). The implication is that there are two kinds of projectivism: “spreading projectivism” (which entails error) and “gilding projectivism” (which does not).⁴ An alternative view is that there is just one kind of Humean projectivism—that “spreading” and “gilding” and “staining” are all metaphorical

¹ Edward Craig (2000) also argues for the compatibility of projectivist and realist interpretations of Hume, but, unlike Sainsbury, Craig thinks that Hume can be interpreted as both regarding *causality*.

² Mackie tends to prefer the term “objectification.” It is clear, however, that he means to capture a kind of projectivism. In his 1980 book he provides a typical description of Humean projectivism (along with the *Enquiry* quote given above), and twice refers to “this projection or objectification” (72). For discussion of Mackie's view of objectification, see Joyce (forthcoming a).

³ A.W. Price (1992) also distinguishes the nihilistic (error-theoretic) form of projectivism from other forms deserving the name: Simon Blackburn's “reductive projectivism” and Richard Wollheim's “genetic projectivism.”

⁴ These are not Sainsbury's labels.

synonyms—and that this projectivism and the metaphors used to describe it are simply neutral on the question of error. One might then hold that in the case of necessary connection the projectivism happens to be coupled with error but in the case of morals it is not. The fact that Sainsbury carefully distinguishes “spreading” from “gilding or staining” indicates that this is not his view.

It may help clarify matters if we begin to break possible projectivisms down into subtheses. (This will be done initially in a rough and ready way, later in a more rigorous manner.) Let’s do it for the particular case of the idea of necessary connection.

1. We experience necessary connection as an objective feature of the world.
2. This experience has its origin in some non-perceptual faculty; in particular, upon observing a regularity in nature we form an *expectation* that brings about the experience described in 1.
3. In fact, necessary connections do not exist in the world.
4. When we utter sentences of the form “X is necessarily connected to Y” we are misdescribing the world; we are in error.

If there is a kind of projectivism that entails an error theory, then it must look something like 1-4. This, I take it, would capture Sainsbury’s “spreading projectivism.” I will call this “nihilistic projectivism.” What of the putative non-nihilistic projectivisms?

In order to render projectivism compatible with noncognitivism, we must delete 4 (or the analog of 4). According to classic moral noncognitivism, when we utter the sentence “X is morally wrong” we are not *describing* the world at all, and therefore cannot be *misdescribing* it. Blackburn, in particular, is keen to emphasize that everyday moral language is not *in error*, despite its projectivist foundation. (If there is an error anywhere, it is the blunder of philosophers who misdescribe the metaphysical commitments of moral discourse.) According to the noncognitivist advocate of projectivism, although our experience may be as of objective moral facts (as in 1), our moral language does not perform the function of expressing the belief that these moral facts obtain; rather (according to an expressivist version of noncognitivism), the function of moral utterances is to express the emotions that give rise to the experiences. (More on this later.)

But although deleting 4 results in a projectivism compatible with noncognitivism, it does not yield a projectivism compatible with realism. In order to achieve the latter, we need also to delete 3. This, I take it, would (*mutatis mutandis*) capture Sainsbury’s “gilding projectivism”—the projectivism he associates with Hume’s stance on morality.

Let us be clear. First, the conjunction of 1 and 2 is neutral between realism and anti-realism; although compatible with realism, the conjunction of 1 and 2 does not entail realism. Nor does it entail either 3 or 4. Second, the conjunction of 1, 2 and 3 entails anti-realism, but is neutral between cognitivism and noncognitivism; although compatible with noncognitivism, the conjunction of 1, 2 and 3 does not entail noncognitivism. Nor does it entail 4. Third, the conjunction of 1, 2, 3 and 4 entails anti-realism in general and

an error theory in particular. Note that no combination of subtheses *entails* realism, and no combination *entails* noncognitivism.⁵

Let us pause to think more carefully about the relation between moral projectivism and noncognitivism. First we had better alter our example to the moral case, along the same lines as we did for causal projectivism:

1. We experience moral wrongness (e.g.) as an objective feature of the world.
2. This experience has its origin in some non-perceptual faculty; in particular, upon observing certain actions and characters (etc.) we have an affective attitude (e.g., the emotion of disapproval) that brings about the experience described in 1.
3. In fact, moral wrongness does not exist in the world.
4. When we utter sentences of the form “X is morally wrong” we are misdescribing the world; we are in error.

I have claimed that nothing here entails noncognitivism, but I can imagine someone tempted to take issue with this. “Surely,” the complaint would go, “if 1 and 2 are true (and we can throw in 3 for good measure) then what lies behind moral experience is an emotion: disapproval. If one then makes public one’s moral judgment—via an utterance of the sentence ‘X is morally wrong’—this sentence thus expresses the disapproval. But the thesis that moral utterances express emotions just *is* noncognitivism. So in fact 4 is incompatible with 1-3; in place of 4 we should have ‘*Therefore*, when we utter sentences of the form “X is morally wrong” we are expressing our emotions.’”

Such an objection would be based on misunderstanding. The metaethical debate between the cognitivist and the noncognitivist does not concern what kinds of mental states *cause* moral judgments; it concerns the linguistic function of moral judgments (whether they are assertions, or commands, or interjections, etc.).⁶ For S’s utterance U to *express* mental state M (in the sense relevant to the metaethical debate) it is neither necessary nor sufficient that U is caused by M. What concerns us here is the sufficiency condition. From the fact that an emotion is causally active in the generation of moral judgment it does not follow that the moral judgment *expresses* that emotion—not, at least, in the relevant sense of “express.” Suppose my child desperately wants a bike for his birthday, and I am inclined to buy him one because I love him and want him to be happy. This love is a central component in the causal chain leading up to me uttering the sentence “I intend to buy you a bike.” And yet this utterance is a straightforward assertion: It has truth conditions and functions linguistically to express a belief (the belief that I intend to buy him a bike). The fact that an emotion has caused the utterance should not distract us from being all-out cognitivists about this utterance. Thus, although

⁵ For no-frills discussion of the relation between realism, noncognitivism, and the error-theoretic stance, see my entry for “moral anti-realism” in the *Stanford Encyclopedia of Philosophy* (Joyce 2007).

⁶ Psychologists sometimes use the term “moral emotivism” to denote the theory that emotional faculties play a central role in the causal generation of moral judgment. Although philosophers are trained to think of emotivism as a kind of noncognitivism, it is clear that in this case the taxonomy doesn’t apply; psychologists are not using “emotivism” in its metaethical sense. See Joyce 2008 for analysis.

subtheses 1 and 2 add up to the claim that affective attitudes (e.g., disapproval) are causally generative in moral judgment—perhaps even necessary for moral judgment—1 and 2 (even with 3 thrown in for good measure) fall short of entailing noncognitivism. 1-3 together *are silent* on the function of public moral utterances.⁷

There is, nevertheless, a tendency in some quarters to think that projectivism entails noncognitivism. Perhaps this is based on the misunderstanding just diagnosed. Or perhaps there is yet another kind of projectivism possible—one that really does entail noncognitivism. Consider these statements by Nick Zangwill, both describing Blackburn’s view:

“Projectivism” is the view that the disputed judgements express non-cognitive mental states, such as emotions, desires, habits, or expectations; but the projectivist also holds that such non-cognitive states are *spread* or *projected* onto the genuine facts and states of affairs. So we come to speak and think as if there were an extra layer of properties in the world. (1992: 161)

According to Simon Blackburn,...“projectivism”...is the view that moral judgements express attitudes (approval, disapproval, liking or disliking, for example), which we “project” or “spread” onto the world. (1990: 583)

Note that both characterizations render projectivism a double-barreled thesis. The second component seems familiar: One could plausibly see the idea of “non-cognitive states...spread or projected onto genuine facts and states of affairs” as a pithy summary of 1 and 2. But the first component is one that we haven’t yet encountered as a subthesis of moral projectivism. This first component explicitly packs the noncognitivist case into the definition of projectivism: “disputed judgments express non-cognitive states, such as emotions, desires [etc.]” This, of course, excludes the possibility of projectivism being compatible with either realism or the error theory.

We now potentially have four kinds of moral projectivism to deal with, to which we can, for the sake of convenience, give the following labels:

⁷ The necessity condition also fails. A speech act (such as an assertion, or an apology, or a promise, or an interjection, or a command) may express a type of mental state (such as a belief, or regret, or a commitment) without the speaker having that mental state. One need reflect only on the phenomenon of *insincerity* to see this. An insincere act of promising still succeeds in being a promise (unlike, say, an overtly sarcastic promise utterance, which is not a promise at all), and, as such, an insincere promise still expresses a commitment on the speaker’s behalf—a commitment that the speaker actually does not have at the time of utterance. Similarly, an insincere assertion (a lie) expresses a belief that the speaker does not have at the time of utterance. But if speaker S can, through uttering U, express mental state M *while S does not have M* (and perhaps has not had M), then the relation between M and U cannot be a causal one. Rather, the relation is a complex one concerning the entrenched linguistic conventions understood by both speaker and audience. If both speaker and audience take it that acts of promising express commitment—if this understanding is a necessary prerequisite to being granted competence with the concept *promising*—then a successful act of promising does express commitment, irrespective of whether the speaker as a matter of fact is committed. (I seem to have said this many times before; in the unlikely event that anyone has been paying attention, I apologize for the repetition.)

- Minimal projectivism: The conjunction of 1 and 2.
- Metaphysical projectivism: The conjunction of 1, 2, and 3. (I call this “metaphysical” because the addition of subthesis 3 adds a metaphysical claim that otherwise isn’t present.)
- Nihilistic projectivism: The conjunction of 1, 2, 3, and 4.
- Noncognitivist projectivism: The ‘double-barreled’ thesis just described, which is minimal projectivism (or possibly metaphysical projectivism) conjoined with expressivist noncognitivism.

| | thesis 1 | thesis 2 | thesis 3 | thesis 4 | expressivism |
|-----------------------------|----------|----------|----------|----------|--------------|
| minimal projectivism | √ | √ | - | - | - |
| metaphysical projectivism | √ | √ | √ | - | - |
| nihilistic projectivism | √ | √ | √ | √ | x |
| noncognitivist projectivism | √ | √ | ? | x | √ |

[NB: The last two columns exclude each other, hence the “x” is to be interpreted as a *denial* (as opposed to simply the absence of endorsement).]

One form of projectivism entails noncognitivism, one excludes noncognitivism, and two are silent on the matter. I do not propose to adjudicate among these theoretical options; it is possible that different kinds of projectivism—even different kinds of *moral* projectivism—are better suited than others for different theoretic purposes. However, I would like to express my misgivings about the noncognitivist projectivism described by Zangwill, for it seems to me to contain an inherent tension. Start by focusing on the second barrel of the first passage: “we come to speak and think as if there were an extra layer of properties in the world”—a comment that jibes with remarks made by Descartes, Galileo, Hume, and the rest. But what is it to “speak and think” as if the world were a certain way? The natural answer is that speaking as if the world were a certain way amounts to *asserting* that it is that way, and thinking as if the world were a certain way amounts to *believing* that it is that way. But asserting and believing are the hallmarks of a *cognitivist* attitude. In other words, the second barrel of noncognitivist projectivism (as stated by Zangwill) seems to presuppose the misfiring of the first barrel. Granted, the tension here is not flat out inconsistency. It is possible to “speak and think” as if the world were a certain way without asserting and believing that it is. A paradigm example of this phenomenon would be acting. The actor utters (and, presumably, thinks) the sentence “Thou art a scholar” without asserting or believing the proposition. So it is possible to reconcile the first and second barrels of this form of moral projectivism, but one will have to tell a special story about speaking-without-asserting and thinking-without-believing in order to do so.

A similar issue arises when we give consideration to whether minimal projectivism entails cognitivism. (Indeed, to the extent that the second barrel of Zangwill’s noncognitivist projectivism is a rough statement of minimal projectivism (or possibly metaphysical projectivism), it’s the very same issue.) One might be tempted to assume that subthesis 1 alone entails cognitivism, for 1 consists of a claim about the nature of

moral experience, which, one might think, implies something about the subject's *beliefs*. Doesn't subthesis 1 amount to the claim that we *believe* moral wrongness to be an objective feature of the world? No, it doesn't. Let us distinguish between moral experience and moral judgment. There is a clear sense in which a stick in water is experienced as bent, but the savvy observer does not judge it (believe it) to be bent. One can imagine a world where minimal projectivism is true but subjects are more or less aware of the fact and are not fooled. Perhaps the inhabitants of this world are all moral error theorists but they continue to have experiences as of an objective moral realm, which they treat as a kind of unavoidable mirage. Or perhaps the inhabitants of this world have located some facts of a subjective (e.g., constructivist?) nature, which they judge worthy referents of their moral vocabulary. These people find that they cannot help but experience these facts as objective, but they know that are not; they do not treat the moral realm *per se* as illusory, but they treat the objective pretensions of that realm as an ineluctable illusion. So subthesis 1—even coupled with 2 to form minimal projectivism—does not entail cognitivism. Nevertheless, it seems fair to say that *generally speaking* how people think and speak about the world goes along with their experience of the world. If people are experiencing the world as containing objective moral properties—as subthesis 1 declares—then it is natural to suspect that people will *believe* the world to contain such properties, and that their language will reflect these beliefs by consisting of *assertions* that the world instantiates such properties. As before, the burden seems to fall on the noncognitivist to tell a special story about how ordinary persons have a critical distance from their moral experience—how they possess a certain kind of sophistication—such that they do not take their experience at face value.

In other words, it seems as if moral cognitivism is the 'natural partner' of minimal projectivism—the default assumption—and that if one wants to endorse minimal projectivism while supporting moral noncognitivism, then one has some explaining to do. The important thing to note is that the projectivist cannot have it both ways: She cannot maintain both (A) that our moral experience, our tendency to “spread” our emotions onto reality, really *fools* us, and (B) that in making a moral judgment we are doing no more than expressing our emotions. What it is for our experience to “fool us” is, presumably, for our *beliefs* go along with how things seem (in which case *assertion* comes naturally along too); but if we are in the business of believing (and asserting), then it is not the case that we are doing no more than expressing our emotions (even if it is true that emotional episodes play a central causal role in the production of our moral experience); we are, rather, squarely in the province of cognitivism.

Getting straight on what kind of moral projectivism is under discussion on any given occasion is of the utmost importance, for it is a truism that before we can embark on investigating whether projectivism is true, we need first decide the content of the thesis whose truth we are scrutinizing. Different data will bear on whether different versions of projectivism are true. In order to investigate the truth of noncognitivist projectivism, for example, we would have to engage with the debate over cognitivism versus noncognitivism. But this debate can be ignored if we are investigating other kinds of projectivism. Alternatively, in order to confirm or disconfirm metaphysical projectivism

we would need to establish to our satisfaction either that there are or are not objective moral facts; whereas if our interest is in minimal projectivism this does not matter.

In what follows I want to focus on subtheses 1 and 2, which together comprise what I have labeled “minimal projectivism.” Despite my calling it this, I am officially agnostic as to whether the conjunction of these two theses really ought to be categorized as a type of “projectivism” or whether it should, rather, be thought of as the common heart of all other kinds of moral projectivism. Perhaps instead of being presented as a delineation of different kinds of moral projectivism, the preceding discussion would be better interpreted of as an exposé of the *confusion* surrounding the idea. Perhaps instead of living with lots of different kinds of moral projectivism, we would do better if we decided on a single unified theory. Whether that all-purpose moral projectivism would be equivalent to what I have labeled “minimal projectivism,” or whether those two subtheses should ultimately be categorized as necessary but not sufficient conditions for moral projectivism, is not something I aim to decide here. There are not, after all, any *facts* about what “projectivism” denotes to which we can appeal to settle these questions; the matter is a pragmatic one, concerning which notion of moral projectivism will serve our theoretic purposes most usefully. This noted, I will, if only for the sake of brevity, continue to refer to this conjunction of subtheses as “minimal projectivism.”

2. Turning a Philosophical Metaphor into an Empirical Hypothesis

It is remarkable how rarely, in all the discussions of moral projectivism over the years, proper attention has been given to the fact that the theory is generally presented and thought about in *metaphorical* terms (e.g., “gilding or staining”).⁸ Even the appellation “projectivism” is metaphorical, for nobody thinks that when a person projects her anger onto her experience of events (say), this emotion literally flies forth from her brain and laminates the world. (Slogan: Projecting emotions is not like projectile vomit.) But how do we determine whether a metaphor is adequate, especially since (it is usually accepted) all metaphors are by definition false? Evidently, projectivism is a theory in need of translation into literal terms before it can be properly assessed. Drawing attention to the conjunction of subtheses 1 and 2 is an attempt to accomplish this.

Another striking feature of moral projectivism that has never, to my knowledge, been properly appreciated is that, to the extent that we can detect something literal lying behind the traditional metaphors, it seems reasonably clear that we are dealing with a thesis that is, either entirely or in part, *empirical*. Once we get past the metaphorical level, we see that projectivism concerns a claim about the nature of moral experience (subthesis 1)—which appears to be an empirical matter—and a claim about the genealogy of that experience (subthesis 2)—which also appears to be an empirical matter. Perhaps, as I say, we should decide that there is more to moral projectivism than just these two

⁸ A notable exception is D’Arms and Jacobson 2006. Simon Blackburn is also no doubt aware of the metaphorical status of references to “projection,” yet (in my opinion) he has done little to replace the metaphor with a precise literal hypothesis. On at least one occasion he confesses that “projectivism” is not an entirely happy term for the position he has so frequently advocated (Blackburn 1995: 36).

claims—so perhaps projectivism will turn out to be not *entirely* an empirical matter—but at the very least it is significant to recognize that a substantive empirical inquiry is an important necessary component of any serious attempt to assess the truth of moral projectivism.

Let us discuss these two subtheses in turn. Doing so requires that we come up with better labels. I will call subthesis 1 “the phenomenological thesis” and subthesis 2 “the causal thesis.”

2.1. *The Phenomenological Thesis*

1. We experience moral wrongness (e.g.) as an objective feature of the world.

Many metaethicists accept the phenomenological thesis. In the debate between the moral realist and her opponents, it is often taken for granted by both sides that the moral anti-realist faces a burden of proof, in as much as it *seems* to us that moral judgments track objective qualities. Moral realists often argue that this represents some kind of burden of proof that the anti-realist must overcome; they argue for moral realism on the basis of the combination of the phenomenological thesis with a methodological principle of epistemic conservatism. For example, we read this from Jonathan Dancy:

[W]e take moral value to be part of the fabric of the world; taking our experience at face value, we judge it to be the experience of the moral properties of actions and agents in the world. And if we are to work with the presumption that the world is the way our experience represents it to us as being, we should take it in the absence of contrary considerations that actions and agents do have the sorts of moral properties we experience in them. This is an argument about the nature of moral experience, which moves from that nature to the probable nature of the world. (1986: 172)

And this from David Brink:

We begin as (tacit) cognitivists and realists about ethics....We are *led* to some form of antirealism (if we are) only because we come to regard the moral realist’s commitments as untenable, say, because of the apparently occult nature of moral facts or because of the apparent lack of a well developed methodology in ethics....Moral Realism should be our metaethical starting point, and we should give it up only if it does involve unacceptable metaphysical and epistemological commitments. (1989: 23-4)

The soundness of this burden-of-proof argument isn’t relevant to our present purposes; I am just noting the endorsement of the phenomenological thesis inherent in this popular form of argument.⁹ (For the most developed version of this argument for moral realism, see Huemer 2005.¹⁰)

⁹ For skepticism about the burden-of-proof argument, see Kirchin 2003; Loeb 2007.

¹⁰ I offer some criticisms of Huemer’s view in Joyce forthcoming b.

The phenomenological thesis is also employed as a premise in arguments favoring moral anti-realism. John Mackie argued that not only is our moral experience as of objective values, but that this objectivism is embedded “in the meanings of moral terms” (1977: 31)—that the assumption that moral values exist objectively “has been incorporated in the basic, conventional, meanings of moral terms” (1977: 35). Mackie goes on to argue that this aspect of morality is in fact not satisfied by the world, and hence he advocates moral skepticism. Again, it is not my intention to evaluate this argument, but rather to note the central role that the phenomenological thesis plays in it.

Nevertheless, for all the widespread support enjoyed by the phenomenological thesis, it has never been properly subject to empirical scrutiny.¹¹ Perhaps the reason for this is that the thesis is seriously unclear on several dimensions, all of which would need to be settled before testing could be undertaken. (I doubt, however, that this typically *is* the reason, since many moral philosophers seem happy to endorse the thesis without worrying about, or attempting to settle—or even, apparently, noticing—the lack of clarity.) There are three conspicuous places where the phenomenological thesis needs refining. First, what is it to experience morality as *objective*? Second, what is it to *experience* morality as objective? Third, what is it to experience *morality* as objective? I do not propose to attempt to settle these questions here, but rather to identify what would need to be settled before anybody should pass judgment one way or the other on the phenomenological thesis. What follows are but preliminary notes.

2.1.1. *Objectivity*:

First, we must get clearer on what kind of *objectivity* is relevant to the phenomenological thesis, for this term is used in different ways by different philosophers. Michael Smith uses “objectivity” to refer to the possibility that moral questions have a correct answer upon which open-minded and clear-thinking agents will converge (1994: 5-6). Crispin Wright associates objectivity with whatever plays a wide (as opposed to narrow) cosmological role (1992).¹² Michael Dummett, by contrast, prefers to argue that sentences of a certain kind are objective if and only if we think of them as determinately true or false though we nevertheless know of no method representing either a proof or a disproof (i.e., the sentences are potentially ‘recognition transcendent’) (1978; 1993). “Objectivity” is often associated with some notion of *mind-independence*, though the matter is far from straightforward, since there are numerous kinds of mind-

¹¹ Those who have come closest are Nichols and Folds-Bennett 2003, and Goodwin and Darley 2007. One might also reasonably claim that the extensive empirical research program concerning the moral/conventional distinction (in developmental psychology, clinical psychology, and cross-cultural studies) has bearing on the phenomenological thesis, to the extent that judgments concerning moral transgressions are taken to be those that (inter alia) hold irrespective of any authoritative decree, which is one way of understanding objectivity. (A reasonable starting point for this large literature is Nucci 2001, Smetana 1993, and Turiel et al. 1987.) However, even these interesting studies do not target the hypothesis that we *experience* morality as objective (as oppose to believe that it is). In my opinion, Goodwin and Darley also employ a misguided notion of *objectivity*.

¹² A subject matter has wide cosmological role if the kinds of things with which it deals figure in a variety of explanatory contexts—specifically, if they explain things other than (or other than via) our judgments concerning them.

(in)dependence relations possible. (Cars, for example, are generally classified as concrete, mind-independent entities, despite the fact that they were designed and built by and for creatures with minds.) We can contrast *existential* mind-independence (X would exist even if no minds existed) with *conceptual* mind-independence (the concept X can be adequately articulated without making reference to any mental entities). For example, if one were to hold that the correct analysis of the concept *moral goodness* is something of the form “*whatever an observer with qualities Q would approve of in circumstances C,*” this would make moral goodness existentially mind-independent (since its instantiation wouldn’t depend on the existence of any such observers) but conceptually mind-dependent (since *approval* is a psychological category ineliminable from the explication). In the case of morality, there is the possibility of a further kind of *practical* objectivity: namely, that moral imperatives have a distinct kind of categorical authority: Maurice Mandelbaum writes that our feeling of being bound by a moral obligation “appears as being independent of preference,...as an ‘objective’ demand” (Mandelbaum 1955: 50).

One response to this abundance of non-equivalent notions of *objectivity* is to judge that what is called for is further discrimination of different varieties of moral projectivism, depending on which distinct kind of objectivity is built into the phenomenological thesis (and consequently into the causal thesis). But I think this would be a profligate and implausible response. It seems unlikely, on the face of it, that just anything that has been given the moniker “objectivity” by philosophers—however legitimately for their local purposes—can be plugged into the phenomenological thesis while still yielding a recognizably *projectivist* theory. It is more plausible that the intuitions lying behind projectivism will be best captured by homing in on a particular kind of objectivity, or a cluster of related kinds of objectivity. It is even possible that upon further reflection we may prefer to eliminate the word “objectivity” altogether in favor of something more unequivocal.

What seems common to all brands of projectivism is that something-or-other is experienced as ‘out there,’ existing, or having certain qualities, antecedently and independently of the subject. Let us postpone the question of how we can *experience* something as ‘out there,’ and just focus on the ‘out-there-ness’ itself. We tend to think of cats, rocks, tables, planets, relative size, chemical constitution, and duration as features of the world, independently of our act of perceiving them. Even if a person has directly *caused* a particular cat’s existence (via arranging a breeding program, say), there is still a robust sense to be attached to the idea that the person does not ‘constitute’ the cat’s existence in the act of apprehending it. But is there anything that we *don’t* think of in this way? Gideon Rosen (1994) has argued that there is little sense to be made of this dichotomy of objectivity/subjectivity—at least with respect to the ubiquitous role it has traditionally played in philosophical debates.

To be sure, we do have “intuitions” of a sort about when the rhetoric of objectivity is appropriate and when it isn’t. But these intuitions are fragile, and every effort I know to find the principle that underlies them collapses. We *sense* that there is a heady metaphysical thesis at stake in these debates over realism ... [b]ut after a point, when every attempt to say just what the issue is has come up empty, we have no

real choice but to conclude that despite all the wonderful, suggestive imagery, there is ultimately nothing in the neighborhood to discuss. (1994: 279)

Consider the emotion of sadness, which is clearly mind-dependent in a perfectly trivial sense.¹³ And yet for any given agent there are many instances of sadness (nearly all of them) of which she is not the author: Nearly all episodes of sadness, for any person, are items ‘out there,’ subjects of discovery, not things she invents or creates in the act of perceiving them. Thus Rosen would doubt that even in the case of sadness have we succeeded in “abrogating the right to think of these facts as robustly real constituents of a mind-independent order” (293). He describes several attempts to frame the distinction—several ways of understanding what it might mean for something to have a ‘less-than-objective’ ontological status—and each he rejects due to the persistent availability of the ‘anthropological perspective’: For any putatively ‘subjective’ phenomenon, Rosen will imagine an anthropologist investigating it, and he will observe that even though the phenomenon may ultimately supervene on psychological states (e.g., pretty much any phenomenon that is the topic of any of the social sciences), the anthropologist nevertheless is, from her own perspective, engaged in the study of a robustly real part of the natural world order.

My purposes here do not require that Rosen’s arguments be countered (though I have briefly critically discussed them in Joyce 2007, and see footnote 15 below); it suffices to note that what we are seeking in trying to clarify the phenomenological thesis of projectivism is something considerably more modest than defending the broad imagery of *objectivity versus subjectivity* that motivates so many philosophical debates. To make clear that we are stipulating a notion just for our local purposes, let me use the term “subject-(in)dependence,” rather than “mind-(in)dependence.” The important thing to note is that we can define *subject-(in)dependence* to be a relativistic notion. Consider a particular episode of sadness: say, Sally’s sadness on Tuesday afternoon. This sadness is subject-dependent relative to Sally; it is subject-independent relative to everyone else. Even for Sally’s boyfriend, who (let’s assume) *caused* the sadness, Sally’s sadness is an item in the world (albeit a psychological item) that is there to be discovered, of which he might be ignorant, of which he is a passive observer, which *could* have occurred without him. There are important questions to answer still—most prominently, what relation precisely does Sally bear to this episode of sadness in virtue of which it is subject-dependent, relative to her?¹⁴—but I don’t propose to pursue them here. I would rather

¹³ By restricting attention to “the *emotion* of sadness,” I hope to put aside tricky (but clearly different) cases involving sad music, sad events, sad faces, etc.

¹⁴ I’m more comfortable saying something about what this relation *doesn’t* consist in. It doesn’t consist in Sally *causing* the sadness, and it doesn’t consist in Sally judging or believing that she is sad. (I’m willing to accept that she may be sad without believing herself to be.) I am tempted to cash it out in terms of *a priori* modal dependence. *This* token episode of sadness (had by Sally on Tuesday afternoon) *could only* have been had *by Sally*. If we imagine a possible world, W, strikingly similar to ours—where there is someone very much like Sally, feeling sadness in very similar circumstances (on Tuesday afternoon, etc.)—but for which we stipulate that (for whatever minimal reason) she is in fact *not* Sally (and does not count even as her modal counterpart, despite the similarities), and nor is anyone else at W, then we would (I suggest)

point out the virtues of taking the path of relativism on this matter. First, by making subject-(in)dependence a relativistic notion we have nullified the significance of the anthropological perspective. If a phenomenon is subject-independent relative to an investigating anthropologist, so what?—it may nevertheless be pertinent to note that (unlike many phenomena) it is subject-dependent relative to some other individual.¹⁵ Second, and more importantly, employing a relativistic notion is all we need. After all, what we intuitively want to capture of the projectivist tendency is the experience a person may have that “I am not the author of this phenomenon; it would carry on the same even unperceived by me.” We do not have to worry about the absolute ‘objective’ status of the phenomenon (or, indeed, whether it is even coherent to think of any such notion of absolute objective status standing in contrast to absolute subjective status); we need concern ourselves only with how the subject experiences it *in relation to herself*.

A third point to reflect upon is the possibility that the kind of ‘out-there-ness’ that the folk employ in their judgments of objectivity is inchoate and in fact resistant to more precise analysis. Although it is natural for a philosopher to seek a more precise understanding of what it is to for a phenomenon to be subject-(in)dependent (relative to an agent), if the everyday notion that figures in people’s thinking is in fact indeterminate, then (A) for the purposes of gauging whether a token judgment is imbued with ‘out-there-ness’ it may not be *necessary* to precisify the relevant notion of subject-independent ‘out-there-ness’ beyond a nebulous and coarse-grained version, (B) it may not be *desirable* to so precisify the notion (since we want to ensure that we’re capturing the folk idea), and (C) it may not even be *possible* to precisify the notion more than we have. Regarding (C), it may turn out that Rosen is entirely correct about the ineliminable confusions lying at the heart of the objective/subjective distinction, but this would not show that people do not *employ* the notion (warts and all). Let us not forget that our task is not to produce a philosophically defensible characterization of *objectivity*—not even as it appears in the phenomenological thesis. Our task is to identify what notion the folk are utilizing, sufficient for us to distinguish those circumstances where they employ it from those circumstances where they do not. If in fact the folk are deeply confused, then a deeply confused notion is the one we should be isolating.¹⁶

conclude *a priori* that *this* token episode of sadness (gesturing to the actual Sally’s actual sadness) does not exist at W. I confess, though, that I am not at all confident that this thinking will give the intuitively correct output across all cases we might want to consider.

¹⁵ Incidentally, from this relativistic notion we could then build an absolute one: A phenomenon is Subject-Independent in the absolute sense (note the upper case) iff there is no perspective relative to which it is subject-dependent. This seems to be something Rosen overlooks. It may well be that for any ‘subjective’ phenomenon we can invoke the anthropological perspective (thus, he thinks, casting the objective/subjective distinction into disarray), but the reverse does not hold. It is not the case that for any ‘objective’ phenomenon (say, the chemical constitution of Jupiter) we can with equal ease invoke the ‘subjective’ perspective, from which some mental activity constitutes the facts of the case. For all Rosen’s arguments, we can still distinguish those cases for which discovery-talk and mind-dependence-talk can co-exist from those cases for which discovery-talk is permissible but mind-dependence-talk is wholly misplaced—and this distinction may be of philosophical significance.

¹⁶ Those with reservations that the folk could possibly be employing a deeply confused or inchoate and indeterminate notion might recall how Socrates typically sets out to demonstrate exactly this: that despite

2.1.2. *Experience*:

There is much that remains to be clarified in what has just been said, but already the second disambiguation of the phenomenological thesis cries out for discussion. How, it might be asked, can this kind of ‘out-there-ness’ possibly be the object of *experience*? Surely (the objection goes) what we experience is far more primitive and simple than anything remotely like this? It is, however, highly debatable how meager or rich the content of experience is. Even confining ourselves to *visual perception*, it has been argued that the contents of experience can include relatively thick properties, such as *being caused by*, *being an object*, *being a house (a tree, etc.)*, and *being subject-independent* (see Searle 1983; Siegel 2006a, 2006b). The last is of particular interest here. One view is that visual experience incorporates *expectations* of how something may change relative to the viewer, and these expectations constitute a phenomenality of subject-(in)dependence; the counterfactuals are not merely beliefs formed by the subject on the basis of visual data, but are properly thought of as part of the visual experience itself (see Merleau-Ponty 1945; O’Regan and Noë 2001). It would be a project of much interest to see whether this line of thought could be plausibly extended to moral experience.

However, it is not necessary to our present purposes to undertake anything so controversial; we do not need to support the view that subject-independence can be experienced *by the senses*. It suffices if subject-independence can figure in mental states—whether these states be perceptual, perceptual based, or otherwise—and there is something *that it is like* to have these mental states. Suppose that the mental states in question are just common-or-garden *beliefs*; there is still a strong case to be made that there is something *that it is like* to have such beliefs occurrently. (See Flanagan 1992; Goldman 1993; Peacocke 1999; Horgan and Tienson 2002; Kriegel 2003; Pitt 2004.)¹⁷ (This option may be unavailable to the noncognitivist projectivist, who may be inclined to doubt the existence of moral beliefs entirely.)

It is also worth noting that the term “experience” is often used in more liberal ways that may have little to do with any state for which there is a fact about what it is like to have it. Consider: “She experienced the fall of Paris in 1940,” “He experienced the bad weather as a personal slight,” “He experienced his mother as overbearing and critical,” “She experienced a great deal of opposition to her project,” “The stock market experienced a slump,” “New Orleans experienced heavy rainfall.” Clearly, there is much variation among these uses; I shall not attempt to classify or analyze them. The point is that there is enough pliability to the term “experience” that we should not too quickly assume that we know what its appearance in the phenomenological thesis amounts to. Specifically, it may be a gloss for something like “We have strong intuitions that moral

confidently employing a term like “justice” or “knowledge,” his interlocutors in fact do not really have any precise idea what they’re talking about.

¹⁷ For what it’s worth, Hume did not think that anything like ‘subject-independence’ was part of the content of sensory experience: “[A]s to the *independency* of our perceptions on ourselves, this can never be an object of the senses” ([1740, book 1, part 4, section 2] 1978: 191).

wrongness (e.g.) is an objective feature of the world.” Perhaps such intuitions have some kind of ‘what-it’s-likeness’ to them; *perhaps they do not*. But even if lacking phenomenal character, sufficiently spontaneous and entrenched intuitions may serve to underwrite the first subthesis of moral projectivism.

Consider, for example, the notion of projection that is often employed in psychopathology, such as when a subject is said to project his feelings as a defense mechanism. Perhaps the example from above, of a person experiencing his mother as critical and overbearing, would be an apt illustration to use. Let us say that his mother is not actually critical and overbearing at all (subthesis 3 of projectivism), but that the subject’s feelings in this respect are caused by his own sense of inadequacy (subthesis 2). Now, although I’ve just used the words “experience” and “feelings” with regard to the subject’s attitude towards his mother’s (supposed) personality, does the projectivist case here really depend on these states having a ‘what-it’s-likeness’ to them? I shouldn’t think so. Perhaps the subject simply *believes* his mother to be like this, and perhaps (contra the views mentioned above) these beliefs have no quality that can be legitimately called “phenomenal character.” This in itself would not undermine the psychopathological diagnosis that the subject is projecting his sense of inadequacy in his dealings with his mother. Similarly, it has been shown that persons suffering from forms of social phobia have a variety of distorted beliefs about social interactions, including assigning a high probability to the proposition that some social gaffe will be committed (Newmark et al. 1973). It seems plausible to say that such persons are projecting their fears onto their view of possible events. Yet the plausibility of this claim does not hinge on the mental act of assigning a high probability to certain events having a phenomenal character.¹⁸

In light of these last comments, perhaps I have named the phenomenological thesis poorly. Yet my use of the term reflects a liberal attitude in metaethics in general (or so it seems to me): Philosophers often refer to “moral phenomenology” meaning “how morality *seems*,” without discussing or even assuming that this “seems” has any phenomenal character in the sense that philosophers of mind intend the phrase. We all know that there are uses of “seems” that do not presuppose phenomenality—e.g., “It seems that dinosaurs went extinct 60 million years ago” (see Tolhurst 1998). I do not, for

¹⁸ I am not claiming that such non-phenomenal ‘experience’ *must* take the form of belief. I should like to maintain the earlier distinction between experience and judgment: One can experience something *as X* while judging that it is *not X*. The introduction of a more fine-grained framework that would accommodate this does not seem objectionable. Let us further consider the phobic, though we’ll change the example to an arachnophobe. Suppose therapy leads the arachnophobe to understand his problem; he comes to realize (all things considered) that the spiders he encounters pose no threat. (We’ll assume he doesn’t live in Australia!) Yet, when he comes upon a daddy long-legs in the bathtub, he finds himself once more in the grip of the thought that the spider is (in some possibly inchoate sense) dangerous. We might choose to accord this ‘thought’ some phenomenal quality (and of course for the phobic this thought is also accompanied by anxiety, which surely does have a phenomenal flavor to it), but doing so does not seem compulsory. Even so, it seems desirable to distinguish the phobic’s thought from a straightforward *belief*. Arguably, the phobic has ceased genuinely to *believe* that the spider is dangerous; he just can’t help entertaining the thought. (For some discussion of the role of thoughts and beliefs in phobias, see Joyce 2000.) Maintaining some logical space between non-phenomenal ‘experience’ and belief also helps to make this way of explicating the phenomenological thesis available to the noncognitivist projectivist (who generally denies the existence of moral *beliefs*).

example, think that the quotes by Dancy and Brink given earlier indicate an intention on their parts to commit to a strong and literal sense of moral phenomenal character.

Even supposing that we *are* talking about the phenomenal character of moral judgments, it is important to note that we are not presupposing that there is something *distinct* about moral phenomenality. It has been noted that moral phenomenology is an approach that presupposes that there is something *peculiar* about the phenomenal quality of morality, such that if there is not, the whole pursuit becomes spurious (see Sinnott-Armstrong 2008; Kriegel 2008). Investigating the phenomenological thesis of projectivism does not engage us in that approach. We are interested in whether moral judgments have a quality of subject-independence. Perhaps judgments about many other things have this quality too (judgments about cats, rocks, tables, planets, relative size, chemical constitution, and duration). Perhaps there is nothing special about the kind of subject-independence that is attributed to morality (assuming that it is). Indeed, one obvious method for investigating whether moral judgments are imbued with subject-independence is to look for similarities—perhaps even exact matches—with other kinds of subject-independent judgment. The phenomenological thesis asserts a simple predication. To think that it implies that there is some special phenomenological ‘signature’ of morality is like thinking that someone who declares that crocodiles are green is committed to there being some distinctive greenness peculiar to crocodiles alone.

2.1.3. *Morality:*

Let me now turn to the third disambiguation of the phenomenological thesis: What does it mean to say that we experience *morality* in such-and-such a manner? There are two kinds of clarification that one would ideally like to see made. First, there is disagreement over how we are to demarcate the moral from the non-moral realm. Is moral normativity necessarily distinct from prudential normativity (as Kant thought)? Are moral imperatives those that one is willing to universalize (as R.M. Hare thought)? Must moral norms concern interpersonal relations (as Kurt Baier thought)? Philosophers argue about such things, and to the extent that these disputes remain unsettled, so too does the domain of the moral. But even if we were to suppose that such worries could be resolved, we would face a second kind of indeterminacy about morality: its tremendous variation. We must distinguish moral decisions from moral judgments (Sinnott-Armstrong 2008); moral judgments of value from moral judgments of duty; moral judgments applied to oneself from moral judgments about others; direct moral judgments from removed moral judgments (Mandelbaum 1955: ch.2-3), first-order from higher-order moral judgments (Horgan and Timmons 2008); moral judgments involving thin evaluative concepts (*good, bad, right, wrong, etc.*) from moral judgments involving thick evaluative concepts (*heroic, sleazebag, wimpy, fair, humiliating, etc.*); and so on.

I have attempted to address the first kind of problem elsewhere (see, especially, Joyce 2006, ch.2), so won't rehearse that thinking again here. Even without settling such disputes, however, it might suffice for our present purposes if we observe that pretty much all parties will agree to certain *paradigms* of moral judgments. When an ordinary person responds to a documentary on Nazi war crimes with the utterance “Those *evil*

bastards!” we will all agree that a moral judgment has occurred, even if we disagree on what qualities were present that warranted this verdict.¹⁹ (We will also agree on *foils* of moral judgments: Nobody is arguing that someone who utters “Taking the bishop with your rook is the best move” has made a moral judgment.) Having a substantial number of paradigms in hand should suffice to test the phenomenological thesis. Finding that the phenomenological thesis holds true of all such paradigms would not, of course, allow us to conclude that it holds of *all* moral judgments, but it would at least be a substantial and interesting start.

We come to a similar conclusion when giving consideration to the second problem. If we were to discover that the phenomenological thesis holds true of, say, direct first-order judgments about one’s own moral duties, we should certainly not conclude that it will also hold true of second-order judgments about moral values. It is entirely conceivable that some but not all of the items on this inventory of moral types will satisfy the phenomenological thesis. But does the phenomenological thesis (or any relevant disambiguation of the phenomenological thesis) really purport to embody a claim about *all* moral experiences? A couple of paragraphs ago I dismissed the supposition that in investigating the phenomenological thesis we are seeking something *distinct* about moral experience; now the question is whether we are seeking something *common* to all moral experience. It is natural to suppose we might be; it is natural to read the phenomenological thesis as a universal generalization. This is in fact something I intend to resist, but I will postpone the matter until after I have discussed the causal thesis.

A more comprehensive paper would now present ideas on how the phenomenological thesis should be tested. That, however, is not my purpose on this occasion. I am satisfied to call attention to the fact that the thesis does amount to an empirical claim, and the only reason one may have for assuming that it cannot (in principle) be tested using scientific methods is thinking that there is something vague, vacuous, ambiguous or incoherent about the thesis. I hope that the preceding comments have gone some way to answering those harboring any of the latter worries: I have tried to show how the thesis can be disambiguated, and I would be surprised if anyone were to think that, so clarified, something incoherent remains buried in the thesis.²⁰ To those persuaded that we now have an empirically testable hypothesis under consideration but who ask “OK, but *how*?” I reply “Good question; let’s try to think of a good answer.”

It is possible that an adequate investigation of the phenomenological thesis will at some point involve an examination of something deserving the name “intuitions”—and if this is so it is vital to avoid the pitfall of assuming that one’s own intuitions—honed by years of metaethics and dripping with theoretic prejudice—should stand in for those of everyone else. If we do attempt to collect a sample of others’ intuitions, it is equally vital that we don’t do so in a shoddy manner (e.g., questionnaires to one’s *Ethics 101*

¹⁹ I say “ordinary person” to exclude certain philosophers, who may hold all sorts of wacky views. Consider what Bishop Berkeley took himself to be saying when he uttered “There is a tree in the quad.”

²⁰ Perhaps certain forms of *objectivity* have been written off as incoherent (see, e.g., Rosen 1994), but, as was noted earlier, the phenomenological thesis doesn’t require that any general concept of *objectivity* ultimately makes sense.

undergraduates at the end of class). Data must be collected in an intellectually responsible manner, complying with the customary scientific standards concerning such things as sample size, control groups, replicability, randomization, correcting for framing effects, and so on.

2.2. *The Causal Thesis*

2. This experience—of morality as an objective feature of the world—has its origin in some non-perceptual faculty; in particular, upon observing certain actions and characters (etc.) we have an *affective attitude* (e.g., the emotion of disapproval) that brings about the experience described in 1.

Investigating the phenomenological thesis promises to be very challenging. But on the assumptions that it can be done satisfactorily (which is, of course, a big assumption), and that the hypothesis receives confirmation, then testing the causal thesis promises to be somewhat more straightforward. Let us suppose that our experience with exploring the phenomenological thesis has left us with a test, or series of tests, that we can apply to subjects in order to gauge their score on an ‘objectivity scale’ with various kinds of experience. (This is almost certainly an idealization of anything we can reasonably hope for, but let us allow ourselves to speak in idealized terms at this preliminary stage.) Testing the causal thesis is a matter of ascertaining whether certain factors causally influence a subject’s performance on this ‘objectivity scale’—in particular, whether emotional arousal has a causal impact.

It seems reasonable to assume that both elements of the causal connection that we wish to investigate—the subject’s level of affective arousal and his/her score on the ‘objectivity scale’—are continuous phenomena. One can, for example, be emotionally aroused not at all, a little, a fair amount, a great deal. We might expect something similar regarding the strength of objectivity with which a person’s moral experiences are imbued. This being so, evidence for a causal connection can be gained via manipulating the hypothesized causal antecedent and observing proportional change in the hypothesized causal consequent. Arousing (certain?) emotions should ‘ramp up’ the subject’s tendency to imbue her moral judgments with objectivity; dampening emotions should be accompanied by a reduced experience of objectivity. (Naturally, standard procedures of randomization, etc., should be enforced.) Of course, this sounds all very easy in principle; no doubt designing adequate experimental protocols will be a far more complicated exercise. One reason I claimed that this might be “more straightforward” than testing the phenomenological thesis (assuming, of course, that the latter thesis has already been tested) is that we know various ways of manipulating subjects’ affective attitudes. Studies in the psychology literature that involve arousing certain emotions in subjects (both openly or surreptitiously) are too numerous to require citing.

Regarding both the phenomenological thesis and the causal, it would be naïve to think that there is any one test that might provide confirmation. In both cases, what we should

be seeking is experimental ‘triangulation,’ whereby we come at the target hypothesis from numerous experimental directions.

3. Moral Projectivism: The General and the Particular

My principal claim is that confirmation of both the aforementioned subtheses would amount to an empirical confirmation of minimal projectivism. But would it be a confirmation of moral projectivism *simpliciter*? That depends on two things. The first we have already discussed: There are conceptions of moral projectivism that require the satisfaction of further subtheses. If our interest lies in one of these other non-minimal versions of projectivism then we should, of course, still be highly interested in the empirical prospects of minimal projectivism, for its confirmation would count as the confirmation of *a necessary part* of our preferred theory. The second complicating factor—the one I earlier postponed the discussion of—is that it’s not clear *how many* token episodes of moral judgment the two subtheses need hold true of before we can legitimately speak of minimal moral projectivism *as a general thesis* holding true. Let me explain.

It seems to me fair to assume that, once basic determinacy and coherence have been accorded to the thesis of minimal moral projectivism, most people will agree that it holds *sometimes*. To the extent that a psychopathological notion of “projecting one’s emotions” is present in vernacular conversation (“He thought that everyone was criticizing him, but really he was just projecting his own insecurities”), it seems plausible to assume that we’re generally comfortable with the idea that *sometimes* moral judgments are the result of individuals projecting their emotions onto their experience of social interaction. Yet one can accept this without thereby counting oneself an advocate of *moral projectivism*. So it seems that a reasonable question to ask is “How frequently would the minimal projectivist account of token moral judgments have to hold before we would claim that Minimal Moral Projectivism is in general true?” (I will now use upper case to indicate the general thesis.)

It seems doubtful that the answer should be “Always.” Think, by comparison, of projectivism about color (bearing in mind Hume’s apparent like treatment of color and morality). The color projectivist need not claim that *every* color judgment is the product of an episode of perceptual projection. If I inform you that my screensaver is predominantly the same color as the sky on a clear day, then you can—without ever laying eyes on my computer—make the judgment that my screensaver is predominantly light blue. One might balk at calling this a “color judgment,” but I have no qualms in that respect. You have the concept *blue*. You come to believe that a particular item (my screensaver pattern) falls within the extension of the predicate “...is blue.” You may then assert the sentence “The screensaver is blue” and thereby say something true.²¹ That sounds like a color judgment to me.

²¹ At least: something that has as good a claim to being true as if you were to assert the same sentence on the basis of visual acquaintance with my screensaver.

There are many differences between making color judgments on the basis of visual apprehension and on the basis of inference. I guess one obvious difference might be sheer *frequency*: Inferential color judgments seem fairly unusual—at least if we’re talking about ones that are explicitly represented in deliberation. This observation might lead one to say that Color Projectivism will be true so long as the projectivist story holds good of *most* color judgments. The fact that projectivism doesn’t hold true of inferential color judgments might be dismissed as statistically insignificant.

But this in fact doesn’t seem to me like the correct way of thinking about the difference. The more salient difference between the two kinds of color judgment is that there is a kind of logical asymmetry between them. The inferential color judgment is parasitic on the perceptual color judgment, in the sense that if someone has *never* made a perceptual color judgment (i.e., has never experienced color), doubt arises as to whether she can even make an inferential color judgment. Consider a person totally colorblind from birth. Such a person can learn that the sky lies within the extension of the predicate “...is blue,” and therefore when it is put to him that my screensaver lies within the extension of the same color predicate, it is a simple matter for him to come to the conclusion that asserting the sentence “The screensaver is blue” will meet with agreement. But does such a person really understand what he’s saying? Does he have the concept *blue* at all? Does he have any beliefs about blue things? I do not need to argue that the answer to these questions is ultimately definitively negative; it is enough for my purposes to note that there is at least a temptation to answer them in the negative. (See Peacocke 1983; Tye 1999.)

My point is that there is a way of understanding the asymmetry between inferential color judgments and perceptual color judgments that is not statistical. Even if *most* color judgments were inferential, there would, I suggest, still be this temptation to treat the perceptual color judgments as somehow privileged. One may, then, claim that what it takes for Color Projectivism to be true is for the projectivist story to hold true of all episodes of color judgment *in the privileged class*. (Or, I suppose, one might say that what is required is that it holds true of *most* episodes of color judgment in the privileged class—thereby mixing in something statistical.)

If this sounds along the right lines, then the same strategy should be available to the *moral* projectivist. Can one make a moral judgment without any act of emotional projection involved? “Sure you can,” says the Minimal Moral Projectivist. The Minimal Moral Projectivist might claim that this happens frequently, or even usually. What makes him nevertheless a Minimal Moral Projectivist is the conviction that there is a privileged category of moral judgment and the minimal projectivist story is true of all (most of?) the members of that class. Suppose there is some kind of impairment that one might suffer—an imaginary impairment will suffice—that leaves us doubting whether the sufferer really has any moral concepts. The sufferer might be savvy enough to catch on to the socially appropriate sentences to utter—she might know that stealing, promise-breaking, and pedophilia fall within the extension of the predicate “...is morally wrong”—and indeed in our conversations with her we may not even realize that anything is amiss. (We might be similarly fooled by a well-trained but completely colorblind person giving a competent

lecture on Impressionism.) But when we discover that the person does in fact have this deficiency—that there is a kind of mental state that she is incapable of having and has never had—we grow doubtful as to whether she really understands what she’s saying, whether she has the concept *morally wrong* at all, whether she really has any moral beliefs.²² If this were so, then the natural thought is that there is an asymmetry in the kinds of moral judgments made by the *unimpaired* persons: that those made in the absence of the mental state(s) in question are parasitic upon those made in its presence. The Minimal Moral Projectivist can then limit his claim to those judgments in the privileged class, irrespective of their statistical frequency.

The question of whether individuals blind from birth have color concepts seems to be an *a priori* one: It is a matter for philosophers to haggle over. It is not so clear that the issue of whether individuals with various kinds of psychological impairment have moral concepts must proceed in an entirely *a priori* matter. Of course, it *might* be that ruminations *au fauteuil* serve to settle the matter; I said above that even an imaginary kind of impairment might be sufficient to ground our conviction that there exists this kind of asymmetry relation. But, on the other hand, it may be that we don’t have very strong intuitions on the matter, and that it is only after a course of empirical inquiry that we feel confident in coming to this conclusion. For example, suppose one hears the following:

Fred suffered brain trauma as a child that left him utterly incapable of feeling empathy. But he often gives normal answers on questionnaires concerning morally-loaded vignettes. For example, he assents to the question ‘Is promise-breaking morally wrong?’ Does Fred have the concept *moral wrongness*?

An uninitiated respondent might not have a strong view on the matter. However, it is possible that when we attend to actual cases of impairment, and carefully examine the subjects’ responses in a variety of domains (perhaps employing a number of experimental methods), we will acquire a body of data that will lead us to conclude that sufferers of this deficiency lack moral concepts. The obvious cases to look at in this respect are psychopaths and people suffering from various kinds of localized brain damage resulting in what has been dubbed “acquired sociopathy” (see Tranel 1994; Bechera et al. 2000; Ciaramelli et al. 2007).

At the risk of annoying the reader by once more stating what I’m *not* doing in this paper: I don’t propose to argue that these kinds of subjects lack moral concepts. My whole point is that this may be a conclusion that we come to only after a careful examination of empirical evidence—much of which may not even be yet gathered. But I will mention that there is already some suggestive data that may point us in this direction. Consider psychopaths. Psychopaths can certainly linguistically respond in an appropriate manner to morally loaded vignettes. (They don’t ask “What does this word ‘right’ mean?” They don’t apply the word “right” to utterly inappropriate things, like days of the

²² The noncognitivist, of course, in a sense denies these things across the board, even for ordinary unimpaired persons. I take it, though, that with a bit of hedging and rewording, the present point about an asymmetry could be expressed in terms amenable even to a noncognitivist.

week or inanimate objects.) And thus they can, at least superficially, demonstrate basic competence with moral terms (as can a blind person competently use color vocabulary.) However, psychopaths can also be found to use moral vocabulary in extremely confusing ways in unguarded moments. One such individual, incarcerated for theft, when asked if he had ever committed a violent offense, replied “No, but I once had to kill someone” (Hare 1993: 125). This is not an isolated slip (see Kennett and Fine 2008 for more examples of this sort of linguistic infelicity in psychopaths). Furthermore, upon more careful examination we discover some strikingly unusual aspects of the psychopath’s deeper grasp of morality. Both children and adults with psychopathic tendencies fail to grasp the distinction between moral and conventional norms—a cross-cultural trait that usually emerges at the age of about three. (For references and discussion, see Blair et al. 2005: 57-8.)²³ Experiments also reveal that psychopaths struggle to process certain linguistic information, especially that which is emotionally salient: Whereas normal persons process emotional words faster than neutral words, for psychopaths there is no appreciable difference (Williamson et al. 1991; see also Blair 2005: 59-62 for further references). Psychopaths lack some of the affective input into linguistic processing, and thus, it may well be argued, suffer from a lack of proper understanding of the associated concepts; arguably, they do not qualify as genuinely knowing what a term like “moral goodness” even means. In his classic study of psychopaths in the mid-20th century, Hervey Cleckley explicitly likened psychopathy to colorblindness: The psychopath cannot comprehend “goodness, evil, love, horror, and humour. ... It is as though he were colour-blind, despite his sharp intelligence, to this aspect of human existence. ... He can repeat the words and say glibly that he understands, and there is no way for him to realize that he does not understand” (1941, p. 90).²⁴ After reviewing a number of sources of evidence, Jeanette Kennett and Cordelia Fine conclude that “a growing body of evidence, ... such as their poor performance on the moral-conventional distinction task and their incompetence in the use of evaluative language, suggests that psychopaths deviate so significantly from the folk that it is reasonable, *on empirical grounds*, to conclude that they do not have mastery of the relevant moral concepts” (2008: 219; my italics).

The case may be more complicated than that of the colorblind person’s apparent failure to grasp color concepts, since it seems to involve an extra logical step. The first step is to argue by direct analogy with the colorblindness case: If a person has never experienced the emotion of guilt, say, then they cannot really have the concept *guilt*. We might repeat this step for a number of different emotion/affect concepts. The additional step is to argue that grasp of these emotion/affect concepts is a necessary condition for being granted competence with the moral concepts. I am not claiming that either step can

²³ Some have expressed doubts about aspects of the moral/conventional distinction (see Kelly et al. 2007; Kelly and Stich 2007), but their skepticism does not extend to casting into doubt the evidence that there exists a substantial performance divergence in this respect in individuals manifesting the psychopathic profile.

²⁴ Kennett (2002) argues that psychopaths lack the concept *duty* due to their impaired understanding of ends and reasons. Smith (1994: ch.3) argues that having certain motivations in favor of compliance is necessary for mastery of moral concepts. He draws a direct analogy with how things stand in the case of a colorblind person’s grasp of color concepts.

be settled exclusively by *a posteriori* inquiry; I am observing that either step may be bolstered and informed by empirical input. A blunt presentation of the question “Fred has no capacity to feel guilt; does he have the concept *guilt*?” may not prompt confident responses. But an examination of the constellation of impairments that follow from this emotional incapacity may help settle one’s views. Of course, ascertaining which other incapacities truly *follow from* the impairment in question, and which are merely contingently associated with it in the case of psychopathy, is an extremely delicate matter. It is, however, very clearly an extremely delicate *empirical* matter.

In sum: Empirical investigations can contribute much to our deliberations concerning the psychological prerequisites for moral competency. Any such conclusion can then allow us to identify a privileged class of moral judgments, which opens the possibility of embracing a non-statistical notion of what it takes for Minimal Moral Projectivism to be true: It is true so long as the minimal projectivist account holds of members of the privileged class; it is not a matter of *how many* token episodes of moral judgment the minimal projectivist account accurately describes, but *which* episodes.

4. Conclusion

In this paper I have attempted to clarify a particular version of moral projectivism and have drawn attention to the fact that it is an empirical thesis. Even if there exist other legitimate versions of projectivism requiring the satisfaction of further criteria, I suggest that at their core will lie the same empirical commitments. The minimal version of moral projectivism that I have discussed is neutral between any of the standard metaethical options—and I would claim this as one of its virtues. One might, then, question whether minimal moral projectivism is *metaethically* interesting at all. Despite the fact that my CV bears the title “metaethicist,” I must confess to being unsure about where to draw the lines around the discipline. (I am none too comfortable about even calling it a “discipline.”) I am quite prepared to endorse an Institutional Theory of metaethics: It concerns whatever metaethicists decide it concerns. Even if minimal projectivism does not count as a “metaethical thesis,” the fact that a proposal that has traditionally been thought of as a metaethical thesis turns out not to be one is itself something of which metaethicists should take note. Similarly, it might be claimed that by homing in on a specifically empirical and metaethically neutral version of projectivism, I am stripping the thesis of its *philosophical* interest. I remain blasé, finding any dispute over whether something counts as philosophically interesting extraordinarily philosophically uninteresting.

Thanks to Walter Sinnott-Armstrong, Neil Levy, Shaun Nichols, and two anonymous *ETMP* referees for their comments.

REFERENCES:

Bechara A et al. (2000) Characterization of the decision-making deficit of patients with ventromedial prefrontal cortex lesions. *Brain* 123:2189-2202

Blackburn S (1993) *Essays in quasi-realism*. Oxford University Press, Oxford

Blackburn S (1995) The flight to reality. In: Hursthouse R et al. (eds) *Virtues and reasons*. Clarendon Press, Oxford, pp 35-56

Blackburn S (1998) *Ruling passions*. Oxford University Press, Oxford

Blair J et al. (2005) *The psychopath*. Blackwell, Oxford

Brink D (1989) *Moral realism and the foundations of ethics*. Cambridge University Press, Cambridge

Ciaramelli E et al. (2007) Selective deficit in personal moral judgment following damage to ventromedial prefrontal cortex. *Soc Cogn Affect Neurosci*: 2:84-92

Cleckley H (1941) *The mask of sanity: An attempt to reinterpret the so-called psychopathic personality*. CV Mosby Company, St Louis, MO

Craig E (2000) Hume on causality: Projectivist and realist? In: Read R, Richman K (eds) *The new Hume debate*. Routledge, London, pp 113-121

Dancy J (1986) Two conceptions of moral realism. *Proc Aristot Soc*, suppl. 60:167-87

D'Arms J, Jacobson D (2006) Sensibility theory and projectivism. In: Copp D (ed) *The Oxford handbook of ethical theory*. Oxford University Press, Oxford, pp 186-218

Darwall S (2000) Normativity and projection in Hobbes's *Leviathan*. *Philos Rev* 109:313-47

Descartes R (1970/1641) Letter to Abbé de Launey. In: Kenny A (ed and trans) *Descartes: Philosophical letters*. Clarendon Press, Oxford, p 109

Dummett M (1978) *Truth and other enigmas*. Duckworth, London

Dummett M (1993) *The seas of language*. Oxford University Press, Oxford

Flanagan O (1992) *Consciousness reconsidered*. MIT Press, Cambridge, MA

Galileo G (1960/1623) *The assayer*. In: Drake S, O'Malley CD (trans) *The controversy on the comets of 1618*. University of Pennsylvania Press, Philadelphia, PA, pp 151-336

Goldman A (1993) *The psychology of folk psychology*. *Behav Brain Sci* 16:15-28

Goodwin G, Darley J (2007) *The psychology of meta-ethics: Exploring objectivism*. *Cognition* 106:1339-1366.

Hare R (1993) *Without conscience: The disturbing world of the psychopaths among us*. Pocket Books, New York

Horgan T, Tienson J (2002) *The intentionality of phenomenology and the phenomenology of intentionality*. In: Chalmers D (ed) *Philosophy of mind: Classical and contemporary readings*. Oxford University Press, New York, pp 520-33

Horgan T, Timmons M (2008) *What does moral phenomenology tell us about moral objectivity?* *Soc Philos Policy* 25:267-300

Huemer M (2005) *Ethical intuitionism*. Palgrave Macmillan, New York

Hume D (1978/1740) *A treatise of human nature*. Selby-Bigge LA (ed) Clarendon Press, Oxford

Hume D (1983/1751) *An enquiry concerning the principles of morals*. Hackett Publishing Company, Cambridge, MA

Joyce R (2000) *Rational fear of monsters*. *Br J Aesthet* 40:209-224

Joyce R (2006) *The evolution of morality*. MIT Press, Cambridge, MA

Joyce R (2007) *Moral anti-realism*. In: *Stanford encyclopedia of philosophy*. Available via <http://plato.stanford.edu/entries/moral-anti-realism/>

Joyce R (2008) *What neuroscience can (and cannot) contribute to metaethics*. In: Sinnott-Armstrong W (ed) *Moral psychology, vol 3: The neuroscience of morality: Emotion, brain disorders, and development*. MIT Press, Cambridge, MA, pp 371-394

Joyce R (forthcoming a) *Patterns of objectification*. In: Joyce R, Kirchin S (eds) *A world without values: Essays on John Mackie's moral error theory*. Springer Press

- Joyce R (forthcoming b) The skeptic's tale. *Philos Phenomenol Res*
- Kelly D et al. (2007) Harm, affect and the moral/conventional distinction. *Mind Lang* 22:117-131
- Kelly D, Stich S (2007) Two theories about the cognitive architecture underlying morality. In: Carruthers P et al. (eds) *Innateness and the structure of the mind: Foundations and the future*. Oxford University Press, New York, pp 348-367
- Kennett J (2002) Autism, empathy and moral agency. *Philos Q* 52:340-57
- Kennett J, Fine C (2008) Internalism and the evidence from psychopaths and "acquired sociopaths" (plus comments and reply). In: Sinnott-Armstrong W (ed) *Moral psychology, vol 3: The neuroscience of morality: Emotion, brain disorders, and development*. MIT Press, Cambridge, MA, pp 173-225
- Kirchin S (2003) Ethical phenomenology and metaethics. *Ethical Theory Moral Pract* 6:241-264
- Kriegel U (2003) Consciousness as sensory quality and as implicit self-awareness. *Phenomenol Cogn Sci* 2:1-26
- Kriegel U (2008) Moral phenomenology: Foundational issues. *Phenomenol Cogn Sci* 7:1-19
- Loeb D (2007) The argument from moral experience. *Ethical Theory Moral Pract* 10:469-484
- Mackie JL (1977) *Ethics: Inventing right and wrong*. Penguin Books, New York
- Mackie JL (1982) *The miracle of theism*. Clarendon Press, Oxford
- Mandelbaum M (1955) *The phenomenology of moral experience*. Free Press, Glencoe, IL
- Merleau-Ponty M (1945) *Phénoménologie de la perception*. Gallimard, Paris
- Newmark C et al. (1973) Endorsement of Ellis' irrational beliefs as a function of psychopathology. *J Clin Psychol* 29:300-302
- Nichols S, Folds-Bennett T (2003) Are children moral objectivists? *Cognition* 90:23-32
- Nucci L (2001) *Education in the moral domain*. Cambridge University Press, Cambridge

O'Regan K, Noë A (2001) What it is like to see: A sensorimotor theory of perceptual experience. *Synthese* 129:79-103

Peacocke C (1983) *Sense and content*. Oxford University Press, Oxford

Peacocke C (1999) Conscious attitudes, attention, and self-knowledge. In: Wright C et al. (eds) *Knowing our own minds*. Oxford University Press, Oxford, pp 63-98

Price AW (1992) Three types of projectivism. In: Hopkins J, Savile A (eds) *Psychoanalysis: Mind and art*. Blackwell, Oxford, pp 110-128

Pitt D (2004) The phenomenology of cognition; or what is it like to think that p? *Philos Phenomenol Res* 69:1-36

Rosen G (1994) Objectivity and modern idealism: What is the question? In: Michael M, O'Leary-Hawthorne J (eds) *Philosophy in mind*. Kluwer, Dordrecht, pp 277-319

Sainsbury RM (1998) Projections and relations. *Monist* 81:133-160

Searle J (1983) *Intentionality*. Cambridge University Press, New York

Siegel S (2006a) Which properties are represented in perception? In: Gendler Szabo T, Hawthorne J (eds) *Perceptual experience*. Oxford University Press, Oxford, pp 481-503

Siegel S (2006b) Subject and object in the contents of visual experience. *Philos Rev* 115:355-88

Sinnott-Armstrong W (2008) Is moral phenomenology unified? *Phenomenol Cogn Sci* 7:85-97

Smetana J (1993) Understanding of social rules. In: Bennett M (ed) *The development of social cognition: The child as psychologist*. Guilford Press, New York, pp 111-141

Smith M (1994) *The moral problem*. Oxford University Press, Oxford

Tolhurst W (1998) Seemings. *Am Philos Q* 35:293-302

Tranel D (1994) Acquired sociopathy: The development of sociopathic behavior following focal brain damage. In: Fowles D et al. (eds) *Progress in experimental personality and psychopathology research*, vol 17. Springer, New York, pp 285-311

Turiel E et al. (1987) Morality: Its structure, functions, and vagaries. In: Kagan J, Lamb S (eds) The emergence of morality in young children. University of Chicago Press, Chicago, pp 155-244

Tye M (1999) Phenomenal consciousness: The explanatory gap as a cognitive illusion. *Mind* 108:705-725

Williamson S et al. (1991) Abnormal processing of affective words by psychopaths. *Psychophysiology* 28:260-73

Wright C (1992) Truth and objectivity. Harvard University Press, Cambridge, MA

Zangwill N (1990) Quasi-quasi-realism. *Philos Phenomenol Res* 50:583-94

Zangwill N (1992) Quietism. *Midwest Stud Philos* 17:160-76