Beyond Sidgwick's Paradigm: A Pluralist's Method for Moral Philosophy

Mark Fedyk[1,2,3]
mark.fedyk@cornell.edu

0.0 Preliminary Distinctions

What is moral truth?—Are moral judgments expressions of feelings? Statements of belief? Representations of facts? Apprehensions of a priori principles?—But also, how can philosophers know the *truth* about any such truths about moral truth?

To keep things organized, let's call answers to questions like "Are moral judgments expressions of feelings?" and "Are some moral judgments warranted?" second-order moral truths, because moral judgements like "I should treat my neighbor as I would have myself treated" are first-order truths, if indeed some, or any at all, of these statements are true. This paper is about the methods we use to formulate and justify both first- and second-order moral truths: its aim is to introduce a set of *third-order propositions* that have important implications for contemporary philosophical debates about both first-order and second-order truths.

I will argue, specifically, that a standard approach to analyzing various first- and second-order moral statements rests on a third-order theory of logic and reference that is unable to handle the complexity of the moral behavior of humans. I will then describe and briefly defend a number of methodological ideas, themselves third-order claims, that taken together form an alternative to the first set of third-order claims. However, it is important to stress at the outset that my argument is not exactly that the second set should replace the first; the thesis of this paper is that the second set of propositions should be taken as a set of alternative but nevertheless complementary methodological ideas about how moral philosophy may go about the business of trying to satisfy its epistemic goals. And the benefit of going in for the alternative methodology is that it has the potential to accelerate the pace at which progress is made towards realizing a goal of moral philosophy by increasing the number of first-order and second-order statements that fall within the scope of moral philosophy.

So, my aim hereafter is not really to do metaethics. And although it is now fashionable amongst philosophers to self-identify as doing, for instance, *meta*-metaethics, that is not my intent either. Instead, my intent is straightforwardly methodological. I want to say enough to distinguish one of the most popular analytical methods of modern moral philosophy from a less well-known alternative. In light of this goal, most of what I will have to say hereafter is neutral

[1] Department of Philosophy, Mount Allison University
[2] The Ottawa Hospital Research Institute
[3] Department of Philosophy, Ashoka University

with respect to many of the standard philosophical debates about important first-order and second-order questions. And because methodologies are only justified by how well they do, *in practice*, promoting epistemic goals, I will mostly be unconcerned with producing a conclusive argument for the alternative method. The value of the alternative method for moral philosophy can ultimately be demonstrated only by trying it out, rather than having a third-order and mostly a priori debate about it.

1.0 A Goal of Moral Philosophy

What are first-order moral judgements about? They are about the moral behavior of humans. The aim of many real-life moral judgments is to assess or redirect some pattern of individually or collectively chosen behavior, by judging whether it is morally appropriate or not, and where these moral judgments occur in the context of set of norms and institutions which motivate people to move towards morally appropriate behavior. Accordingly, one of the foremost aims of moral philosophy is to understand whether there are rational criteria upon which any such moral judgments can be, or should be, based.

By saying that moral judgements are about moral behavior, I making a more simple and basic claim than a host of related second-order propositions—for example, that it is correspondence between first-order moral truths and certain properties of behavior that makes the first-order statements true, and so the criteria that should be used to assess first-order statements is evidence of either correspondence or its absence.

And by saying that one of the leading goals of moral philosophy is to discover whether there are rational criteria upon which to evaluate moral judgments, I am not also suggesting that the search for such criteria has proved especially fruitful, or alternatively that we have proof that this search will fail. Rather, the point of placing into our discussion an assumption about one of moral philosophy's goal is that doing this allows us to hereafter focus more directly on the issue of whether there are different *methods*—different sets of third-order propositions—that can facilitate progress towards realizing this goal.

2.0 Truth and Complexity

Is Aristotle short? That is: is this sentence true: "Aristotle is short"? According to a popular view of truth, the sentence is true just in case the referent of the sentence, Aristotle, has the property *short*, and the sentence is false if Aristotle otherwise lacks this property. More importantly, the truth-value of the sentence is determined by a complex calculation according to which atomic lexical objects with root semantic values are "assembled" into the sentence by applying a finite set of recursively-applicable syntactic rules, such that the truth-value of the sentence is a logical consequence of only the application of the syntactic rules and the semantic values of the root lexical objects. Furthermore, the only semantic values that the root lexical objects can take are 'true' and 'false', and so the only truth-values that "Aristotle is short" can take are likewise 'true' and 'false'. The conjunction of a bivalent theory of truth (the claim that the only semantic values are 'true' and false') and a compositional semantic theory (the claim

that the meaning and thus the truth-value of a statement is determined by its syntactic parts plus their root semantic values) are what we can call the standard semantic theory.

What is recognized in the philosophy of science is that this standard semantic theory is, quite literally, an idealization. The root of the idealization is the implicit, but false, metaphysical view that the mereological structures of the world's causal systems are almost always isomorphic to the linguistic structures that can be formed by application of the relevant recursively-applicable syntactic rules and a bivalent theory of truth. In fact, there are two failures of isomorphism here. The first isomorphic failure arises between the semantic values that root lexical items are assigned according to model theory for first-order logic and the real properties and kinds of the world. Here, the "alignment" or "fit" between the root semantic values and real properties and kinds is never so exact as to prevent definitional mismatches, borderline cases, and imprecision. Descriptive frameworks are logically consistent sets of predicates that have some explanatory or inductive value; the forms of "misalignment" between descriptive frameworks and the world explain why it is never the case that we can assign all the predicates of a single descriptive frameworks in a 1-to-1 fashion all the properties of causal systems.

The second failure of isomorphism occurs because neither the syntactic rules of composition that fix the meanings of sentences in first-order logic nor the much richer set of mathematical formula used by various natural and social sciences can be used to make *generally applicable* calculations about how systems will behave as they undergo change. While there are many types of systems in the world, none undergo changes *as if* they follow only the logical and syntactic rules of first-order logic, and nor have we discovered any set of mathematical theorems that tells us how properties will be inherited, emerge, suppressed, balanced, or stabilized over the history of any causal system whatsoever. There are no true generalizations about all systems, because all systems—which include ecologies, salts, the economy of California, and even Aristotle himself—do not have the properties or causal powers that they have as *only* a consequence of how they are composed out of smaller entities and processes. Sodium ions, for example have very different causal powers than salts. And then, what causal powers a particular salt has can depend whether it is dissolved in a solution or remains a solid; many salts are insulators as solids but become highly conductive when dissolved.

The lack of isomorphism between the predicates of a descriptive framework and the world's properties and kinds is the root of many contemporary discussions of vagueness. To see an example of this, let us suppose that Aristotle is 4 feet and 0.1 inch tall. Is it true to assert "Aristotle is 4 feet tall"? No—on the standard semantic theory, this sentence is false. But if Aristotle is 4 feet and 0.1 inches tall, then it is true to assert "Aristotle is short"? Yes—but only if we allow that some truths are vague or imprecise, which the standard semantic theory does not allow. As the familiar sorites arguments show, there is no one property that can be the referent of predicates like "short", and so either the semantic value of the predicate "short" cannot be explained on the standard semantic theory and we have to reject the general applicability of this theory (Braun & Sider 2007), or we have exclude from our vocabulary any potentially vague predicates. The later is untenable, since *all* natural language predicates are

vague. As Paul Teller has shown (Teller 2014), when using natural language predicates, we always face a choice between vagueness (imprecision) and accuracy. The sentence "Aristotle is 4 feet tall" is false on the assumption that in reality Aristotle is 4 feet 0.1 inches tall. But we can inject some vagueness into the sentence in order to preserve its truth: "Aristotle is roughly 4 feet tall". However, if it is protested that "Aristotle is roughly 4 feet tall" is insufficiently precise, then we can remove the source of the sentence's vagueness by turning the sentence into "Aristotle is exactly 4 feet 0.1 inches tall". That sentence is true, but only *ex hypothesi*. In real life, no one and nothing is ever exactly 4 feet 0.1 inches tall, because everything is constantly changing due to shifts of atmospheric pressure, gravity, and in the cases of human beings, such things as even the quality of lunch.

Let us say that what makes something a *complex* system is that its metaphysical structure undergoes changes according to patterns that are not isomorphic with any structure that can be built in a *single* descriptive framework. But what does it mean to "build" a "structure" in a descriptive framework? The standard semantic theory gives an answer to this question. The axioms of a formal language stipulate that there is a set of primitive expressions and syntactic rules for combining the primitive expressions. There will also be a formal model, which consists of a set of properties, a function defined so that its domain is that set of properties, and the value the function takes ('true' or 'false') as it maps these properties onto the primitive expressions. The value of this function then becomes the truth-value of the relevant primitive expressions. Then, any expression in the language counts as a *grammatical* expression only if it is either a primitive expression or a by-product of applying the rules to either some of the primitive expressions or the by-products of prior applications of the rule. Grammatical expressions just are the structures that can be *built* within the descriptive framework, and the usual mathematical, boolean, logical, and syntactic operators are the rules out of which these structures are built. Grammatical expressions are true only if they are composites of lexical primitives that themselves have the value 'true'. So to say that there is an isomorphic relationship between the structures built in a descriptive framework and the metaphysical structure of some natural process is to say lexical items in the descriptive frameworks can be mapped in a 1-to-1 fashion onto all of the properties of the relevant system, and also that the causal behavior of the natural process acts *as if* it obeys, in a law-like fashion, either the combinatorial syntax of a formal language or the mathematical expressions into which the the framework's lexicon can be embedded. Calculations within relevant descriptive framework, then, will be exactly accurate inferences about any possible behavior of the system because the syntactic or mathematical transformations that are possible within the descriptive framework will be in perfect alignment with all of the possible real-world behavior of the system. Consequently, a *complex* system is any system that requires more than a single descriptive framework in order to be described—simply because the system's behavior is too complicated to be captured by any single descriptive framework.

Here is an example to help make things more clear. Mechanical watches are sophisticated machines built out of hundreds of gears, springs, pulleys, and so on. Nevertheless, mechanical watches are built out of only different types of the six simple machines, and they can be

designed and manufactured using only classical mechanics. Are mechanical watches *complex* natural processes? They are. Despite appearances, there is no single description of the physical components and internal behavior of a watch that can also predict, with perfect accuracy, the timing of the watch. Because of that, mechanical watches satisfy our definition of a complex natural system. Even the watchmaker working with the best machined parts and perfect knowledge of the watch "blueprint" will be unable to predict if the watch she is assembling will run fast or slow. That is: she cannot take measures of all of the parts of the watch and then calculate the exact speed at which the watch will run. Part of the reason for this is that watches are not closed systems that are completely buffered from external forces—moving the watch closer or further from the moon will subtly influence its timing. But one of the most germane reasons why watches are complex systems is that a watch, from the moment its movement starts to spin, wears on itself, causing changes in the friction and torque created by its parts. Every watch will wear in a different way as it is affected by all sorts of outside forces (vibration, moisture, gravity, one-off impacts) and that is why no watch is, technically speaking, a natural process that is isomorphic to the conjunction of its blueprint and timing calculations.

At this point someone might object that mechanical wear can be described mathematically in all sorts of different ways, and that it therefore is possible to establish the relevant isomorphisms by expanding the definition of system to include all the different causes of wear. But unfortunately that will not work. At least some of the wear will be the byproduct of phenomena at the quantum level, and, notoriously, it is not possible to consistently integrate quantum mechanics with classical mechanics. So what this shows is that there are at least two descriptive frameworks needed to describe the real-world behavior of a particular watch. There is the classical mechanical timing calculations, which define the watch's "ideal" behavior—and then the quantum mechanical calculations that explain one of the ways that this particular watch is deviating, even if every so slightly, from the ideal. So, while it is true that constructing two different descriptive frameworks to explain the behavior of a watch allows us to know more about it, the fact that two descriptive frameworks are necessary to predict the watches' behavior is itself evidence of the watches' complexity in the relevant technical sense.

Now, given the definition of complex natural processes, nearly *everything* is a complex natural process—because nearly every system that humans have studied has required the formulation of multiple descriptive frameworks that apply at different levels of generalization and with different sets of idealizations and simplifying assumptions. The explanation for this that all descriptive frameworks—whether they be purely mathematical models or mixed linguistic and logical theories—must employ what linguists call generic terms, and which are terms that refer to entire categories or kinds of properties (Leslie 2008). But the kinds themselves, whether they are natural or social in nature, do not have exact boundaries. The idea here is well said by Locke when he writes that "nature, in the constant production of particular beings, makes them not always new and various, but very much alike and of kin one to another: but that the boundaries of the species whereby men sort them, are made by men." (Locke et al. 1860, p.374) Consequently, there is a mismatch—a fuzziness, or vagueness—built into the referential link

between generic terms and kinds. Language is organized by different forces and constraints than those shaping the world's kinds; but both language and kinds are *orderly*, not random. Both can therefore be made to fit together in a coherent fashion, but the fit is rarely if ever perfectly logical in its character.

So, if we take the standard semantic theory literally, it follows that there are no truths: no statement about natural phenomena is both perfectly accurate and perfectly precise. But it has been a near constant refrain in the philosophy of science for almost a half-century now not to expect science to produce descriptive frameworks that are uniformly accurate and precise. All scientific theories involve the idealizations and simplifications—this holds in physics (Cartwright 1983; Cartwright 1997), chemistry (Weisberg 2006), biology and the social sciences (Wimsatt 1974; Cartwright 2016; Cartwright & Hardie 2012; Rodrik 2015). William Wimsatt put the general conclusion this way over 40 years ago: "in biology and the social sciences, there is an obvious plurality of large, small, and middle-range theories and models, which overlap in unclear ways and which usually partially supplement and partially contradict one another in explaining the interaction of phenomena at a number of levels of description and organization." (Wimsatt 1974, p.67) This is all evidence that the standard semantic theory cannot explain how scientific language functions in supporting the inductive and explanatory successes that are plainly observable in many different sciences [cf. (Boyd 2001)].

Now, contradictory descriptive frameworks are only a serious problem if the standard semantic theory is taken literally, because a contradiction between two descriptive frameworks entails that some of the vocabulary of one of the frameworks must be false. But if the standard semantic theory is itself treated as an idealized—a rough and approximate—model about how language and the world are related, then contradictions between descriptive frameworks become less worrisome. Two contradictory descriptive frameworks can be interpreted as different rough approximations of the same system. So, instead of taking the standard semantic theory literally and then using that interpretation as a reason to police descriptive frameworks for contradiction, we can take the standard semantic theory to be an idealization that provides us with reason to discount the importance of ensuring the logical consistency of all of the descriptive frameworks we are prepared to accept, which then frees us to spend our time trying to embellish any number of the many additional epistemic virtues that descriptive frameworks can exemplify, beyond logical consistency.

3.0 Sidgwick's Paradigm

That said, many contemporary moral philosophers spend a great deal of time and energy formulating descriptive frameworks out of sets of first-order and second-order moral claims, and they often argue for or against these frameworks on grounds of their logical consistency or inconsistency with different *prima facie* plausible observations, philosophical intuitions, or any of the background beliefs which constitute the relevant philosopher's common sense. And so far as I know, Sidgwick was amongst the first contemporary moral philosophers to demonstrate the philosophical utility of this methodology. Sidgwick's main concern in *The Methods of Ethics* is to evaluate on a comparative basis different first-order moral theories and also a variety of

second-order theories for deciding which first-order moral statements to accept. For the most part, what Sidgwick calls *methods* are individual descriptive frameworks that are formed out of unique sets of first-order and second-order claims, and which generate rational criteria for evaluating moral judgment and behavior (Sidgwick 1981).

But there is also a "methodology" to Sidgwick's assessment of different ethical methods. This methodology is a set of third-order claims that, early in his book, Sidgwick sets out to use as the framework upon which his subsequent analyses are based. To keep things clear, then, I will refer to Sidgwick's methodology by using the word "paradigm", and so we can say that Sidgwick's paradigm is characterized by these (but not only these) statements:

(1) "I therefore assume as a fundamental postulate of Ethics, that so far as two methods conflict, on or the other of them must be modified or rejected." (6)
(2) "My difficulty begins when I have to choose between the different principles or inferences drawn from them [i.e. the individual discrete methods]. We admit the necessity, when they conflict, of making this choice, and that it is irrational to let sometimes one principle prevail and sometimes another; but the necessity is a painful one. We cannot but hope that all methods may ultimately coincide." (14)
(3) "The aim of Ethics is to systematize and free from error the apparent cognitions that most men have of the rightness or reasonableness of conduct." (77)

Sidgwick is being characteristically careful in his use of "cognitions" in the third quotation. For him, this term functions as a catchall term that refers to all ethical judgments, feelings, habits, intuitions and so on. The term, in other words, refers to any mental activity that induces a person to act in a way that they perceive to be morally right.

Thus, Sidgwick's paradigm is a third-order theory about how philosophers can and should go about addressing both first-order and second-order questions. It says that moral philosophy should systematize moral cognition, where this involves establishing (at least) a logically consistent set of moral principles and associated moral judgments. It is clear that, for Sidgwick, a moral theory must be both internally consistent and as comprehensive as possible: Sidgwick sees the ethicist's task as one of freeing common sense morality of its contradictions while also showing how moral judgments can be based on a small number of basic principles. Crucially, however, this can only be carried out if the first-order and second-order moral statements that interest Sidgwick are interpreted as requiring no choice between precision and accuracy.

As I intimated above, I think Sidgwick's paradigm has evolved to become the predominant methodology used in contemporary Anglo-American philosophical ethics. Certainly, it is easy to find debates over first-order and second-order moral claims that seem to presuppose it. For example, consider the debate between David Copp and both Terry Horgan and Mark Timmons over moral realism. Horgan and Timmons summarize the debate this way:

> Our generic Moral Twin Earth argument is intended to undermine versions of naturalistic moral realism – a type of metaethical view according to which moral terms like 'good' and 'right' refer to natural properties of some sort, and so moral judgments of the form, 'A is good' and 'A is right' are true just in case A has the relevant natural property to which the moral terms refer. This form of moral realism is supposed to be incompatible with versions of moral relativism; if two individuals or groups make what appear to be conflicting moral judgments about some object of moral evaluation, then, assuming they are using moral terms properly, they are engaged in a genuine disagreement and at least one group must be mistaken. (Horgan & Timmons 2000, p.139)

The influence of a literal interpretation of the standard semantic theory is plain to see. Timmons, Horgan, and Copp may be disagreeing about moral statements that went largely unconsidered by Sidgwick in his *Methods*, but it seems very much that they are still hewing to his paradigm. In light of this, we can add a fourth claim to Sidgwick's paradigm:

> (4) The standard semantic theory should be interpreted as a literally accurate picture of how moral language and moral properties are related to one another.

The implication of this claim is that the work of systematizing moral cognition can be carried out by attempting to construct the largest possible set of acceptable first-order and second-order moral statements, and where one (but not the only) factor which determines the acceptability of the relevant statements is whether they are logically consistent with the other members of the same set.

4.0 Theoretical Unity in Moral Philosophy

The desire to build a logically-consistent system of first-order and second-order moral statements or judgments is easy to justify. Any such system can, *inter alia*, provide us with guidance about how to make ethically sound choices in situations we have never encountered before, it can increase philosophical understanding of the ethical project, and it can lead to the discovery of new ethical principles, and so on. Crucially, then, creating a logical system of first-order and second-order moral statements is a way by which rational criteria for evaluating discrete moral claims can be identified—and it is therefore one method by which a fundamental goal of moral philosophy may be realized.

Yet, systematizing the claims of morality does not automatically mean that we should hew to Sidgwick's paradigm. We have, in fact, two choices about how we go about systematizing morality. There is the path taken by applying Sidgwick's paradigm, in which logical consistency is one of the most important criteria used to evaluate the acceptability of first-order or second-order moral claims. Ensuring that the eventual system maintains its logical consistency cannot be the only principle in this paradigm, of course, since at the very least further principles (or evidence of any other kind) will be needed in order to guide decisions about which of two logically inconsistent statements should be allowed into the system. More

importantly, though, is the fact that carrying out this work requires privileging a literal interpretation the standard semantic theory, because assessing the logical consistency of two propositions just is (even if only implicitly) to apply the standard semantic theory to uses of the relevant statements.

The other path is the process of formulating multiple descriptive frameworks for a domain that attempt to cover as much of the phenomena within the domain as possible, but without the requirement that the frameworks be logically consistent with one another. So, we can reserve the world "systematize" for the process of pursuing a single logically-consistent descriptive framework that cover the most phenomena within a domain as possible, and thereby distinguish "systematizing" from its alternative, "organizing".

Now, virtually all sciences are in the business of organizing, not systematizing, and this means that scientists cannot define theoretical unity in terms of maximal logical consistency. It also mean that scientists must augment an appreciation of the value of logical consistency with an appreciation of additional theoretical values, such as novelty of prediction, explanatory breadth, explanatory depth, comprehensiveness, and so on. More important is the idea that the different descriptive frameworks produced by scientific activity may eventually be accepted as scientific knowledge because different frameworks exemplify different theoretical virtues. The mathematical simplicity of one framework can be paired with the comprehensiveness of an alternative qualitative framework, and differences in theoretical virtue like this can compound the depth of understanding conveyed by learning the confederacy of theory formed by collecting together all of the individual descriptive frameworks.

Thomas Kuhn gives us a nice example of what pursuing organization instead of systematization looks like. He notes that even purely technical concepts like "mass", "photovoltaic effect", and "spin" can take on different meanings as these concepts become embedded in the different descriptive frameworks deployed within even a very small discipline, like physics:

> Consider … the quite large and diverse community constituted by all physical scientists. Each member of that group today is taught the laws of, say, quantum mechanics, and most of them employ these laws at some point in their research or teaching. But they do not all learn the same applications of these laws, and they are not therefore all affected in the same ways by changes in quantum mechanical practice. On the road to professional specialization, a few physical scientists encounter only the basic principles of quantum mechanics. Others study in detail the paradigm applications of these principles to chemistry, still others to the physics of the solid state, and so on. What quantum mechanics means to each of them depends on what courses he has had, what texts he has read, and which journals he studies. (Kuhn, Thomas 1970, p.49)

There are few entirely a priori debates in physics about the single true meaning of, for instance, "mass". Physicists, like all other scientists, are comfortable with a degree of descriptive

relativism, but only so long as this relativism is employed in the service of sustaining incremental increases in scientific understanding. And as Kuhn famously argued, conjoining this observed descriptive relativism with the standard semantic theory leads to the conclusion that scientific theories are incommensurable with one another, and that it is therefore not possible to make any judgments about the progress of scientific inquiry writ large. But as I have suggested several times now, another alternative is to not take the standard semantic theory literally.

Indeed, we have no way of knowing if Sidgwick would accept the standard semantic theory if it were presented to him, as his work predates the development of first-order model theory. Still, his method is mostly consistent with the role that the standard semantic theory plays in contemporary analytic philosophical ethics. A large part of book three of Sidgwick's *Methods* is an examination of the descriptive relativism inherent in common-sense moral judgments, and Sidgwick's response to this relativism is to engage in a process of rational reconstruction—his aim, again, is "to systematize and free from error the apparent cognitions that most men have of the rightness or reasonableness of conduct."

Thus, Sidgwick's paradigm has as its methodological end the formation of a single logically-consistent set of first-order and second-order moral claims that, because of the demands of consistency, must all take the same meanings and be expressed in the same vocabulary. The alternative methodology can tolerate logical contradictions and differences in meaning between the statements embedded in different descriptive frameworks—just so long as the relevant statements are interpreted as approximations and thus inherently imprecise claims. Success, according to Sidgwick's paradigm, is a single consistent descriptive framework, while success according the alternative methodology will be a confederacy of different descriptive frameworks that maximize a balance of different theoretical virtues. Because of that, let us call the methodological alternative to Sidgwick's paradigm *descriptive framework pluralism.* The methodological goal of Sidgwick's paradigm is systematization, while the methodological goal of descriptive framework pluralism is organization.

5.0 The Complexity of Moral Behavior

Nothing in our argument so far provides us with a reason to prefer descriptive framework pluralism to Sidgwick's paradigm. Certainly, there is no inference from the fact that the sciences have implemented descriptive framework pluralism to the conclusion that the methods of moral philosophy should, too, hew to descriptive framework pluralism. So in this section I introduce two independent, but parallel, arguments that support the idea that moral philosophy can benefit from adopting pluralism as one professionally acceptable methodology.

Both arguments aim to establish the same conclusion, *viz.* that the moral behavior of humans is complex in the technical sense introduced above. Given as much, the mandate of Sidgwick's paradigm to pursue systematization means that its users in moral philosophy will be able to evaluate much less phenomena than they would be able to if they adopted descriptive framework pluralism. So, relative to the aim of discovering rational criteria for assessing moral

choices and actions, the pluralist's method should do better than Sidgwick's paradigm because it will cover more choices and actions. The argument for descriptive framework pluralism, then, is that employing it will increase the number of both first-order and second-order statements than fall within the purview of moral theory—and that this expansion in the scope of moral theory will be instrumental in the realization of moral philosophy's goal of discovering whether there are rational criteria upon which to evaluate moral behavior, because moral theory will then cover comparatively more possible cases of moral thought and action (i.e., first-order statements) as well as more possible evaluations or explanations (i.e., second-order statements) for the relevant first-order statements.

5.1 Induction from Behavioral Ecology and the Social Sciences

Normative behavior is not the sole provenance of humans. It is an established fact that all living members of the hominidae are able to regulate their behavior by, at least, very simple norms of self-interest and prudence (Tomasello & Call 1997), and the same is probably true of several species of monkey as well (Drayton & Santos 2016). We know very little about the specifics of the cognitive processes that, at the psychological level, represent these norms, and even less about the inferential patterns that link together these representations with patterns of individual and social behavior. It is safest to say only that we know that it is a capacity of the great apes and some monkeys to engage in normative behavior—see, e.g., (Lakshminarayanan & Santos 2008).

But we also know that multiple different descriptive frameworks are necessary in order accurately describe the proximate causes of both individual and social norm-guided behavior in these primate species. Technically, this is because of the difficulty involved in confirming hypotheses about how the the cognitive processes function in non-human primates (MacLean et al. 2014; Leimgruber et al. 2016/5; de Waal 2007), as primate cognition seems to involve the interaction of many different causal processes at differing levels of physiological, neurological, cognitive, and social organization. In that respect, the cognition of nonhuman primates is no different than the cognition of human primates: both are complex enough to impose upon those who would study it the requirement to construct multiple different descriptive frameworks (Marr et al. 2010).

And then, since moral behavior is a kind of normative behavior, it is likely on inductive grounds that moral behavior is similarly complex—it is likely, that is to say, that multiple different descriptive frameworks will need to be constructed in order to sustain any reasonably comprehensive understanding of moral behavior.

Yet, this is inference unlikely to be persuasive to a philosopher who believes, plausibly, that the normative behavior of humans is categorically different than the normative behavior of primates. Note, I do note say that, also plausibly, human normatively behavior is categorically different than the normative behavior of non-human primates *because* it is categorically more complex. That may be true, but it also begs the question of the argument.

This objection, then, can be addressed by the fact that in the social sciences progress with respect to the development of good theories of normative behavior has presupposed descriptive framework pluralism, and this fact supports the relevant inductive inference. The history of social scientific inquiry over the 20th century has not seen the various different social sciences collapse down into a single descriptive framework. Even Gary Becker was happy with the idea that the usefulness of his analytical framework was amplified by logically incommensurable historical and sociological theories (Becker et al. 2012).

To sharpen these observations, allow me to introduce another technical concept: let us say that patterns of normative behavior are comprised out of *dimensions*, which are the actual or possible stable regularities at different levels of organization that are or can be, frequently enough, jointly responsible for causing the different patterns of observable behavior. So, some of the dimensions involved in producing the patterns of behavior which constitutes a vote in the US presidential election will include, *inter alia*, the production of mass-media content, the growth or lack thereof in the economy, the level of engagement and attention that millions of different people are able to hold vis-a-vis complex political issues, and so on. Now, what the history of the social sciences shows us is that there is a many-to-many relationship between the dimensions of any pattern of normative behavior and the descriptive frameworks that are proprietary to different disciplines in the social sciences. Even simple economic behavior is not the sole provenance of economics: sociology, history, and philosophy all have useful things to contribute to our aggregate knowledge of, for instance, the price system. But at the same time, it is certainly not the case that, as Becker of course showed, the descriptive frameworks of economics only apply to situations in which people or firms are making purely fiscal decisions about future actions and tradeoffs.

Now, ethical and moral behavior are *dimensions* of many different patterns of normative behavior. And there is a many-to-many relationship between the dimensions of the relevant social behavior and the different disciplines that study the relevant behavior. On inductive grounds, then, it is very unlikely that the ethical and moral dimensions of human social behavior are the one counterexample to the trend in both behavioral ecology and the social sciences: it is very unlikely, that is to say, that a single descriptive framework will prove sufficient to account for all that we can hope to know about the moral behavior of our species, when both our species and several of our nearest evolutionary neighbors have social behavior that is complex enough to give rise to a many-to-many relationship between the dimensions of the behavior and the descriptive frameworks used to express our aggregate knowledge of the relevant dimensions.

It may be objected here that my argument only concerns the question of whether *describing* moral behavior will require one or many descriptive frameworks—and that the argument does not touch the deeper issue of whether, in order to find a rational basis for *directing* or *guiding* moral behavior, that it will likewise be necessary to construct, evaluate, and consolidate a plurality of descriptive frameworks that are capable of generating normative recommendations. The next section addresses this objection directly.

5.2 Unintended Progress in the Social Sciences

We observed above that one of the aims of moral philosophy is to determine whether there are rational criteria upon which moral judgments can be based. On the basis of this claim, it can be demonstrated that different projects in different social sciences have made what might be termed unintended theoretical progress—because practitioners of the relevant social sciences have, mostly unintentionally, discovered either important first-order moral statements (i.e. descriptions of moral phenomena, or moral prescriptions) or meaningful second-order moral statements (i.e. criteria that can be used to evaluate or explain first-order moral statements). In fact, the thesis that there is a many-to-many relationship between *all* dimensions of normative behavior and the descriptive frameworks constructed by different disciplines entails that disciplines outside of moral philosophy will generate either first-order or second-order moral knowledge. Because of that, the argument that moral behavior is complex in the relevant technical sense can be made by way of a few examples.

Consider first of all the moral credentialing effect (Monin & Miller 2001; Kouchaki 2011). This effect is induced when participants in a psychological study are first offered the opportunity to "morally self-credential" with simple pro-social values by considering a hypothetical moral dilemma that can be resolved by asserting the applicability of the values to the dilemma. Participants who have self-credentialed are latter observed to more frequently violate the pro-social values that they used to self-credential when compared with participants who have not self-credentialed. As a logical matter, the effect cannot be described without employing technical concepts from descriptive frameworks proprietary to moral philosophy, social psychology and statistics. But much more important is that fact that the putative discovery of moral credentialing licenses an investigation into a large number of novel first-order and second-order moral claims. Examples of some claims that could be shown to be either true or false by these investigations include: "people should avoid making all conscious ethical judgments before engaging in moral behavior", "moral rationalism is self-defeating", "teaching people about moral self-credentialing improves their ability to act in accordance with the moral values that they accept", and "we should teach people about moral self-credentialing".

A different by not unrelated example is provided by a study recently completed by Melissa Sands (Sands 2017). Sands found that when randomly selected people observed poor people in public spaces like public park or subway platforms, they were less likely to support taxing the extremely rich than people who had not just seen a poor person. As with the previous study, this result raises a number of important moral and metaethical questions; digging into these questions may involve evaluating moral statements like "it should be illegal to employ actors to impersonate poor people outside of polling stations" and "feelings of disgust lead people to make moral judgments supporting punishment".

It probably does not need to be argued that the Clinton-era reforms to the US welfare system were, amongst other things, a profound change to the moral fabric of American society. Indeed, many of the arguments both for and against the reforms were explicitly moral arguments, and

many of these arguments made predictions that the reforms would either harm or help both welfare recipients and society considered as a whole. Confirming or disconfirming these ethical predictions is not something that can be done from the armchair. It takes different kinds of social-scientific expertise to work through the effects of the reforms on everything from labor productivity, unemployment, physical health, subjective well-being, and so on [cf. (Slack et al. 2007; Ziliak 2009)]. As before, the relevant point is simply that no single descriptive framework, and indeed, no one discipline has descriptive resources sufficient to capture all that we need to know in order to generate an evidence-based answer to the question of whether Clinton's welfare reforms were the right thing to do, morally speaking. And doing the work needed to answer this question will of necessity generate a host of first-order moral statements as well as a host of second-order statements that can be used to justify or controvert some of the first-order moral statements. Some examples include: "women were unfairly harmed by the reforms", "the reforms increased the *de facto* fairness in the US labor market", and "a change governmental policy is justified by a cost-benefit analysis of all the marginal differences it makes".

One of the upshots, then, of the pluralist perspective is that more disciplines than just professional moral philosophy are making contributions towards realizing the goals of moral philosophy. That is why, to a pluralist, ethics is amongst the most interdisciplinary intellectual projects running through the humanities and the social sciences. Conversely, a proponent of Sidgwick's paradigm faces the difficult task of trying to systematize all of the first-order and second-order moral insights generated by work in the social sciences—or, failing that, of providing a convincing argument as to why all these apparent moral insights (it does not matter whether they are true or false) aren't really *genuine* moral insights at all.

5.3 The Relation Between Sidgwick's Paradigm and Descriptive Framework Pluralism

It is tempting to conclude that what we have demonstrated at this point is that Sidgwick's paradigm must be rejected, on the grounds that it cannot accommodate the complexity of moral behavior. Yet, that is not exactly the right conclusion to draw. Instead, what should be rejected is the view that Sidgwick's paradigm is the *only* set of third-order claims that should be used by moral philosophers. As a technical matter, a descriptive framework pluralist can take Sidgwick's paradigm to apply "locally", to one or more first- or second-order descriptive frameworks that are candidates for inclusion in moral philosophy's overall explanatory confederacy. Indeed, a pluralist can and should accept that Sidgwick's paradigm is simply one of several methods for undertaking moral philosophy. A proponent of the Sidgwickian method must deny pluralism, but the pluralist cannot completely reject Sidgwick's paradigm.

Thus, my argument is not that pluralism and Sidgwick's paradigm are inconsistent with one another, and that there is an exclusive choice between the two. Rather, the point is to raise the status of pluralism, and that is as far as my argument can go, because, for the pluralist, Sidgwick's paradigm is a perfectly serviceable method for developing individual descriptive frameworks. It just cannot be that Sidgwick's paradigm is taken to imply that, furthermore, moral philosophy should attempt to construct one single, monolithic descriptive framework.

6.0 Pluralism as a Method for Moral Philosophy

So much, then, for the difference between pluralism and Sidgwick's paradigm. What would it mean to use descriptive framework pluralism as a in moral philosophy? We have already seen that a pluralist will not take the standard semantic theory as a literal representation of how moral language and properties are related. But beyond this, what further third-order provisions would define the pluralist approach to moral philosophy? Here are some suggestions:

1. Treating the standard semantic theory as an idealization should lead moral philosophers to discount, relative to the methodological status quo, the importance of logical consistency and analytical precision in discussions of most moral phenomena. This would make it easier to follow the advice of Bernard Williams, who writes that "our major problem now is actually that we have not too many but too few [ethical ideas], and we need to cherish as many as we can." But we need not follow Williams's suggestion by taking up particularism (Dancy 1983): it can be an aim of moral philosophy to embed various cherished insights into a number of small- and medium-scale descriptive frameworks that, taken together, provide us with a greater ability to apply our insights to equally small- and medium-scale moral problems. So, part of what it means to discount logical consistency is to choose to not worry above all else whether different small- and medium-scale descriptive frameworks contradict one another, and also to choose to reject that idea that all of the aims of moral philosophy must be met by way of the construction of a single, systematic descriptive framework. A pluralist can, instead, use the descriptive framework(s) most appropriate to different aims of moral philosophy, and see the emergence of different definitions for concepts like "harm", "right", "well-being" as an indication of progress, since a pluralist should not expect that any ethical concept, thick or thin, has a single precise definition.

2. Nevertheless, as the prior quote from Kuhn indicates, pluralism is not a form of anything-goes relativism: a many-to-many relationship is not the same as an any-to-any relationship. The continued employment of any descriptive framework in the confederacy of of a discipline's theory is always of question whether the relevant framework continues to have a comparative advantage over any alternative frameworks in helping the discipline realize its epistemic goals. Descriptive framework pluralism is not a pernicious form of relativism.

But the pluralist must rely upon the structure of the world, not logic and a priori reasoning, to ensure consilience is maintained between between the descriptive frameworks employed by a discipline—and this requires that the discipline's practitioners acquire and maintain a "feel" for the phenomena they are studying. Research in the sociology of science has demonstrated that this "feel" is frequently a precondition for the transmissions of scientific knowledge: what one group of scientists know about the behavior of, for instance, a laser can only be transmitted to a new group of scientists if and when the later group has acquired a deep enough practical facility with (or "feel" for) the parts of the laser, the ins-and-outs of its assembly, the oddities of micro-variations in voltage, and so on (Collins 1974; Collins & Pinch 1993; Collins 2001). For

the pluralist, it is the world itself—and our continued openness to its structure—that holds the descriptive frameworks together.

There are two important methodological implication of this point. First of all, in order for some moral philosophers to take up the pluralists method, philosophers must carry out their work in a way that bring them much, much closer to the phenomena that moral philosophy is about. For example, this would require a turn away from discussing moral judgments in a generic mode—such as when moral philosophers talk about the schema of a judgment "x is wrong" in completely decontextualized fashion, and apply the judgment schema only to deeply hypothetical cases. Second, it means that moral philosophy should become more interdisciplinary. Since the dimensions of moral behavior stand in a many-to-many relationship with the disciplines which study the relevant social behavior, the comprehensive evaluation of moral thought and behavior cannot be accomplished by the discipline of moral philosophy alone. Accepting as much leads one to the stance that many different social sciences will be relevant to the project of moral philosophy—as I suggested above, this means ethics may be one of the most thoroughly interdisciplinary intellectual projects.

3. The comparative advantage of professional moral philosophy in the academy cannot, on the pluralist's perspective, rest on training moral philosophers in *only* the techniques of formulating exquisitely subtle a priori definitions and making extremely fine-grained logical analyses. In addition to this, the set of professionally-viable skills for moral philosophy should be broad enough to include the ability to *organize*—that is, to build empirically-realistic "big picture" theories of moral thought and action, which serve to integrate moral insights generated by any of the disciplines which study dimensions of normative behavior. One of the most famous quotes from Wilfrid Sellars puts the relevant idea better than I can:

> The aim of philosophy, abstractly formulated, is to understand how things in the broadest possible sense of the term hang together in the broadest possible sense of the term. [...] To achieve success in philosophy would be, to use a contemporary turn of phrase, to 'know one's way around' ..., not in that unreflective way in which the centipede of the story knew its way around before it faced the question, 'how do I walk?', but in that reflective way which means that no intellectual holds are barred. (Sellars 1963)

A pluralist can happily take up this concern. But the map she makes of how morality "hangs together" will not be a single logically unified topographical representation of the moral domain. Rather, it will be more like a book of maps, each with different scales, keys, projections, and so on. Many pages in this book may be borrowed from different social sciences, and all will deserve entry in the book only because they serve one of moral philosophy's ends.

4. For the methodological pluralist, no one descriptive framework is more fundamental than any other. Even when there are logical relations between conceptual components of two

different descriptive frameworks, the pluralist will nevertheless act only *as if* the standard semantic theory is true, and thus take these logical relations to be, for instance, a simplified representation of how two phenomena described by the frameworks are related to one another in the world. This point is especially relevant to the methodology of analytic moral philosophy, because there is a tendency, most extreme in the Kantian sub-tradition, to prioritize the psychological level of conscious and deliberate practical reasoning over all other levels.

By contrast, a pluralist can treat one or multiple levels of psychological description as important components of an overall confederated theory of the rationality or irrationality of patterns moral behaviors without requiring that the psychological descriptive frameworks do all the work needed to satisfy the goals of moral philosophy. Said differently, the metaphysical structure of normative behavior to a pluralist is nothing at all like the organization of a mechanical watch. Even if there is Kant's "good will", for a pluralist there is no mandate to place a the good will at the very center of moral theory, treating it as a mental mainspring which can move in just the right way all of the other parts of individual choice, action, and society, and by so doing create patterns of truly moral choice and action.

7.0 Conclusion

For most of the first half of the 20th century, philosophers assumed that nature was organized into overlapping systems governed by natural laws, and it was the business of scientific research to discover these laws. Each law was an exact, precise, necessary, and universal causal relationship. Natural laws would be very hard to break, would be extremely resilient against outside interference, and could be observed acting in the exactly the same across many different contexts. It is this picture of nature which support the idea that there can be isomorphisms between the structure of natural world and theories of it that are formulated through applications of the standard semantic theory. But the plausibility of this "systems of systems of laws" picture of the world started to change in the 1970s and 1980s as philosophers, historians, and sociologists of science came into closer contact with scientific practice. As a result of this contact, it soon was discovered that discovering laws of nature is *never* what scientists *ever* do. Instead, and in the language of philosophy of science, scientists are in the business of building and fine-tuning *models* (Giere 2004; Giere 2009)—that is, small-scale and medium-scale descriptive frameworks. All of these descriptive frameworks, both individually and collectively, fall far short of being uniformly accurate and precise representations of natural regularities. Instead, they are a decent—with well-known simplifications, falsifications, idealizations, and omissions—pictures of some of the more interesting dimensions of complex causal systems. Their applicability is highly context sensitive.

There is strong tradition in moral philosophy of conceptualizing ethical principles like natural laws used to be understood: ethical principles are frequently said to be necessary and categorical in their applicability, and that it is a conceptual mistake to think otherwise. But has moral philosophy succeeded in systematizing first- and second-order moral statements around any such principles? It is probably too soon to tell. Still, Bernard Williams surely counts as someone who is sceptical that the systematizing project in moral philosophy has made enough

progress to justify itself and thereby vindicate Sidgwick's paradigm. But Williams also thought, incorrectly, that "science has some chance of being more or less what it seems, a systematised account of how the world really is." (Williams 2011, p.135) So, for him, there is a categorical difference between how moral philosophy can and should arrange public representations of its knowledge and how scientists should do the same. The truth, however, is that it is not longer reasonable to think that science can produce a systematic account of the natural world—in science, we will have to settle for as much organization as turns out to be possible and not any more.

The alternative to Sidgwick's paradigm, then, is not necessarily what Williams believed it to be, namely, an unorganized collection of insights about what is to be done. Moral philosophy can be something quite like the most successful biological and social sciences: an organized but not logically unified collection of small- and medium-sized descriptive frameworks, each of which either allows us to make a moral evaluation of some pattern of behavior, or provides us with guidance about how to evaluate some other moral judgment, and so on. By taking up the pluralist's method, moral philosophy preserves its ability to make progress towards its ends as well as its ability to preserve its cherished insights both. But it must do away with the third-order theory that says that moral philosophy must proceed by constructing a single deductively unified descriptive framework. Moral philosophy can organize without systematizing, even though this alternative is, admittedly, a much messier enterprise. Nevertheless, we should prefer the mess if it allows us to make greater progress towards knowing which, if any, first-order and second-order moral statements are true.

References

Becker, G., Ewald, F. & Harcourt, B.E., 2012. Becker on Ewald on Foucault on Becker: American Neoliberalism and Michel Foucault's 1979 "Birth of Biopolitics" Lectures.

Boyd, R., 2001. Truth Through Thick and Thin. In R. Schantz, ed. *What is Truth?*. Berlin, New York: DE GRUYTER.

Braun, D. & Sider, T., 2007. Vague, So Untrue. *Noûs*, 41(2), pp.133–156.

Cartwright, N., 1983. *How the Laws of Physics Lie* 1 edition., Clarendon Press.

Cartwright, N., 2016. Loose Talk Kills: What's Worrying about Unity of Method. *Philosophy of science*, 83(5), pp.768–778.

Cartwright, N., 1997. Where Do Laws of Nature Come From?*. *Dialectica*, 51(1), pp.65–78.

Cartwright, N. & Hardie, J., 2012. *Evidence-Based Policy: A Practical Guide to Doing It Better* 1 edition., OUP USA.

Collins, H.M., 1974. The TEA Set: Tacit Knowledge and Scientific Networks. *Science studies*, 4(2), pp.165–185.

Collins, H.M., 2001. What is tacit knowledge. *The practice turn in contemporary theory*, pp.107–119.

Collins, H. & Pinch, T., 1993. *The Golem: What Everyone Should Know About Science* Reprint edition., Cambridge University Press.

Dancy, J., 1983. Ethical Particularism and Morally Relevant Properties. *Mind; a quarterly review of psychology and philosophy*, 92(368), pp.530–547.

Drayton, L.A. & Santos, L.R., 2016. A decade of theory of mind research on Cayo Santiago: Insights into rhesus macaque social cognition. *American journal of primatology*, 78(1), pp.106–116.

Giere, R.N., 2009. An agent-based conception of models and scientific representation. *Synthese*, 172(2), pp.269–281.

Giere, R. n., 2004. How Models Are Used to Represent Reality. *Philosophy of science*, 71(5), pp.742–752.

Horgan, T. & Timmons, M., 2000. Copping out on Moral Twin Earth. *Synthese*, 124(1), pp.139–152.

Kouchaki, M., 2011. Vicarious moral licensing: the influence of others' past moral actions on moral behavior. *Journal of personality and social psychology*, 101(4), pp.702–715.

Kuhn, Thomas, 1970. *The Structure of Scientific Revolutions* Second., University of Chicago Press.

Lakshminarayanan, V.R. & Santos, L.R., 2008. Capuchin monkeys are sensitive to others' welfare. *Current biology: CB*, 18(21), pp.R999–1000.

Leimgruber, K.L., Rosati, A.G. & Santos, L.R., 2016/5. Capuchin monkeys punish those who have more. *Evolution and human behavior: official journal of the Human Behavior and Evolution Society*, 37(3), pp.236–244.

Leslie, S.-J., 2008. Generics: Cognition and Acquisition. *The Philosophical review*, 117(1), pp.1–47.

Locke, J., M., A. & O'M., T., 1860. *An Essay concerning Human Understanding. Thirty-second edition. With notes and illustrations of the author, and an analysis of his doctrine of ideas. Also questions on Locke's Essay by A. M. [who also signs himself T. O'M., i.e. Thaddeus O'Mahony.]*, W. Tegg.

MacLean, E.L. et al., 2014. The evolution of self-control. *Proceedings of the National Academy of Sciences of the United States of America*, 111(20), pp.E2140–8.

Marr, D., Poggio, T.A. & Ullman, S., 2010. *Vision: A Computational Investigation into the Human Representation and Processing of Visual Information*, The MIT Press.

Monin, B. & Miller, D.T., 2001. Moral credentials and the expression of prejudice. *Journal of personality and social psychology*, 81(1), pp.33–43.

Rodrik, D., 2015. *Economics Rules: The Rights and Wrongs of the Dismal Science* 1 edition., W. W. Norton & Company.

Sands, M.L., 2017. Exposure to inequality affects support for redistribution. *Proceedings of the National Academy of Sciences of the United States of America*. Available at: http://dx.doi.org/10.1073/pnas.1615010113.

Sellars, W., 1963. *Science, Perception, and Reality*, New York: Humanities Press.

Sidgwick, H., 1981. *The Methods of Ethics* 7th edition., Hackett Publishing Company.

Slack, K.S. et al., 2007. Family Economic Well-Being Following the 1996 Welfare Reform: Trend Data from Five Non-Experimental Panel Studies. *Children and youth services review*, 29(6), pp.698–720.

Teller, P., 2014. Language and the Complexity of the World.

Tomasello, M. & Call, J., 1997. *Primate Cognition*, Oxford University Press.

de Waal, F., 2007. *Chimpanzee Politics: Power and Sex Among Apes*, JHU Press.

Weisberg, M., 2006. Water is Not H2O. In D. Baird, E. Scerri, & L. McIntyre, eds. *Philosophy Of Chemistry*. Boston Studies in the Philosophy of Science. Springer Netherlands, pp. 337–345.

Williams, B., 2011. *Ethics and the Limits of Philosophy*, Taylor & Francis.

Wimsatt, W.C., 1974. Complexity and Organization. In *PSA 1972*. Boston Studies in the Philosophy of Science. Springer Netherlands, pp. 67–86.

Ziliak, J.P. ed., 2009. *Welfare Reform and its Long-Term Consequences for America's Poor*, Cambridge University Press.

Ziliak, J.P. ed., 2009. *Welfare Reform and its Long-Term Consequences for America's Poor*, Cambridge University Press.