The Yale PHILOSOPHY REVIEW

An Undergraduate Publication

ISSUE 5|2009

§

The *Yale Philosophy Review* is an annual journal that showcases original philosophical thought by undergraduate students, worldwide. The goal of the *Review* is to promote philosophical discourse of the highest standard and to bring together a community of young philosophers both in the United States and abroad. Each issue contains a selection of essays on a broad range of topics as well as an interview with a notable contemporary philosopher.

ISSUE V, 2009

For online archive and submissions guidelines, visit <www.yale.edu/ypr>.
Email the editors at yalephilosophy@gmail.com.

EDITORS' NOTE

At times, philosophy can seem very much like an "academic" enterprise. We read arguments, consider objections, and write papers. But the questions we pose do not stay buried in 17th-century manuscripts or dog-eared course packets; they challenge us to answer them and, perhaps, to try to live our lives based on the answers. The purpose of the *Yale Philosophy Review* is to encourage undergraduates, at Yale and elsewhere, to see philosophy as "theirs." In writing about our own ideas and reading the work of peers, we acquire a broader perspective on what it means to participate in the philosophical endeavor: to become active members of a conversation among individuals of widely varying backgrounds, temperaments, and aspirations.

In this issue of the *Review*, Samuel Bagg considers the relationship between morality and scientific developments. What can research into our biological structure teach us, if anything, about the way we *should* act? Fahd Husain presents an original reading of Heraclitus, who helps us think about how our quest for truth can come to terms with an irreducible dimension of ambiguity in the world. Nal Kalchbrenner tackles Kant's views on geometry, raising the question of how our spatiotemporal intuitions can adapt to non-Euclidean geometries. Benjamin Hersh's essay on naturalized epistemology asks: can philosophy really show us the meaning of knowledge, or ought we turn to empirical psychology to show us what we know and how? Finally, Puneet Dhaliwal addresses the relationship between consequentialism and rights. If we focus on the effects of an action when determining its ethical value, must we accept violations of rights in order to produce better overall consequences?

We hope that perusing our fellow undergraduates' thoughts on questions like these will introduce us to fruitful ways of thinking and, perhaps, encourage us to try our own hand. As Professor Hilary Putnam, eminent philosopher at Harvard University, comments in our interview in this issue: "I don't view the truth as lying in the past." We are pleased that the search for truth will continue long into the future, and we are happy to be able to provide undergraduates with a forum in which to start.

We would like to thank the Yale Philosophy Department, the Program in Ethics, Politics, & Economics, and the Undergraduate Organizations Funding Committee for their extremely generous support.

Rachel Bayefsky & Dominic Zarecki,
Editors-in-Chief

The Yale Philosophy Review

Issue V, 2009

CONTENTS

The Fifth Way of Biologicizing Ethics: Science as the Engine of Moral Progress

SAMUEL BAGG
Yale University

What can science say to moral philosophy? Not much, according to most moral philosophers. In a certain way, they are right – the is/ought divide cannot be crossed any more easily now than centuries ago. This paper argues, however, that a scientific investigation of our moral nature reveals the traditional scope of moral philosophy to be far too narrow. Modern moral theories like deontology and consequentialism typically focus entirely on influencing the process of moral reasoning, disregarding the emotional and motivational processing that determines moral output in most cases. Instead of relying on rational rules to guide us through every dilemma, this paper argues that we should turn to a scientifically informed virtue ethics in order to craft better people and more humane societies.

Introduction

"It is those who know little, and not those who know much, who so positively assert that this or that problem will never be solved by science." – Charles Darwin

If we try to follow the conversations of philosophers and scientists, it often seems as if they aren't speaking the same language. Especially when it comes to discussions of morality, there is such indignant anger among both communities at the preposterous contentions of their rivals that there would appear to be enormous moral disagreement between the two camps. But in reality, most scientists and philosophers want basically the same things in the world, and unfortunately, they are prevented from working together because of linguistic confusion and a gulf of understanding that neither side is willing to cross.

In this paper, I attempt to sketch the nature of this divide and some causes of it, and to recommend a way in which scientists and philosophers can come to a reconciliation. In particular, I focus on the history of the discussion of biology's role in moral theory, using Philip Kitcher's *Four Ways of Biologicizing Ethics* as the central

challenge that anyone wishing to incorporate science into moral philosophy must answer. I argue that a practical approach to morality grounded in virtue ethics, especially of the kind advocated by Paul Churchland in *Toward a Cognitive Neurobiology of the Moral Virtues*, holds promise as a legitimate response to this challenge, but that it needs some reconsideration before it can do so. I conclude that a refined practical approach to morality is the "fifth way" of biologicizing ethics, and that moral philosophers would be irresponsible not to incorporate it into their discipline.

Virtue Ethics at the Crossroads of Moral Thought

"We become just by performing just action, temperate by performing temperate actions, brave by performing brave actions." – Aristotle

All moral inquiry can be reduced to a single, fundamental question: how shall I live? It is most natural for modern man to see the answer as a definite rule or set of rules captured by either deontology or consequentialism. Under this conception, moral debate occurs on the fringe of the moral spectrum, using test cases that are often entirely theoretical. This is because the differences between contemporary moral theories don't concern everyday things – most people agree on the answers to such banal problems as whether to steal, lie, or cheat, so it is only in the theoretical domain that modern debates rage on.

There is another approach, though, favored by the ancients,[1] which is not so concerned with hammering out the details of indestructible moral rules. The ancient approach focuses on a daily morality rather than one that seems to live entirely on the periphery of our moral consciousness. According to this conception of morality, the question of how to actually *become* a basically good person in the real world is more pressing than debates over what to do, theoretically, in extreme cases. This attitude I will call the "practical" approach to morality, since its primary focus is achieving a set of pre-existing moral goals, rather than attempting to systematize or ground those goals in some unassailable metaethical sense. Aristotle and other classical authors are the strongest advocates for such an approach, which was generally grounded on the concept of "eudaimonia": the "happiness" of a human being who is "functioning well."

With the onset of modernity, however, the practical approach began to lose favor. Its questions were unspecific, its terms undefined and its answers unsatisfying. What is "eudaimonia," and why is it good? Why is one particular way of functioning better than any other? The ancients were often unclear on these most important of points – they were advocating a certain course of moral learning on the assumption that it was the best without doing the hard work of proving its superiority. With the help of

[1] Including, for example: Aristotle, Plato (under some interpretations), the Cynics, Stoics, and Epicureans.

the Judeo-Christian ascendance, the modern approach to morality began to dominate, frustrated with the superficiality of the ancients' practical response. It raised deeper questions: why behave ethically in the first place? How can I tell for sure whether a specific action is morally correct, especially if I have conflicting intuitions about it? According to this paradigm, represented both by the deontology of Kant and the consequentialism of Bentham, everything which concerned the ancients could be addressed only after a coherent, objective moral code had been worked out. This I will call the "theoretical" approach, because it sees the practical problems of everyday existence as insufficiently interesting and focuses instead on the more fundamental difficulties of systematizing and grounding moral rules that are applicable to every situation, be they entirely common or once-in-a-lifetime dilemmas.

But modernity has come full circle, and the past century has seen the invention of behavioral and human sciences with previously unimagined potential for precision. The tide has begun to turn again, and the "practical" canon has been revived with the advent of the predictive, scientific approach towards morality, which seeks to analyze and explain the ideas of morality as they exist in actual minds and societies, rather than speculate about what their ideal existence might be. In this project, psychologists, sociologists, and evolutionary biologists have schematized the acquisition of what they assume to be moral concepts without defending their version of morality.[2] For example, Lawrence Kohlberg, one of the founders of developmental psychology, represented "moral development" as a six-staged process starting with self-interested concerns and ending with adherence to abstract moral rules; many would say in response that his concept of "development" assumes too much about what sorts of attitudes towards morality are advanced and which are primitive.

Philosophers have been dumbfounded. Have these scientists learned nothing about the problems of moral analysis, some complain, from the last two thousand years of philosophy? The scientists, on the other hand, view the philosophers with a similar disdain – how can they completely reject the insights of science, especially when they have had so long to figure out the answers to these questions and have made so little progress? As a result, scientists and philosophers often refuse to engage, instead willfully and contemptuously talking past one another.

[2] See, for example, Lawrence Kohlberg, "Stage and Sequence: The cognitive-developmental approach to socialization," in *Handbook of Socialization Theory and Research*, ed. D. A. Goslin (Chicago: Rand McNally, 1969), pp. 347-480; John Finley Scott, *Internalization of Norms: A Sociological Theory of Moral Commitment* (Edgewood Cliffs, NJ: Prentice Hall, 1971). They approach the question from different sides of a debate within psychology – is morality a natural developmental process, as Kohlberg holds, or a learned behavioral paradigm, as Scott holds? Both, however, take the content of moral concepts as given in an attempt to show how they develop in humans.

There is one notable exception. Within the discipline of philosophy, there has been a similar, though not entirely parallel, trend back towards practical concerns, which has great promise as the bridge over these troubled waters. The "virtue ethics" of Aristotle – a prime example of the practical tradition of ethical thought – has been revived at the insistence of influential writers like Elizabeth Anscombe and Alasdair MacIntyre.[3] Unlike the scientists, these writers don't usually disregard outright the traditional questions of modern normative and meta-ethics. However, their focus is only "orthogonal" to those disciplines – which have dealt mainly with strictly codified moral rules. The primary aim of virtue ethicists is to show "how we should *be*" all the time "rather than what we should *do*" in specific cases.[4] Virtue ethics concerns itself not only with the kind of tough moral dilemmas addressed by deontological or consequentialist maxims, but with the organization of everyday life. It places emphasis on developing well-trained emotional and behavioral reactions – also known as the virtues – in place of well-developed rational theories.

It might seem that virtue ethics is just begging the question – that is, attempting to get out of the hard question of "what ought we do?" by answering simply: "that which is virtuous." But this criticism is unfair. What separates virtue ethics from most modern moral theory is not a complete neglect of the questions of what we should do, but rather a unique, additional interest in other questions.[5] Virtue ethics doesn't necessarily reject the idea that there are finer points to be worked out in normative ethics, or even that the grounding of morality is an important and troubling challenge. Some virtue ethicists attempt to provide answers to these questions and some don't, but all would agree that definitive, provable answers aren't necessary for moral progress. Instead, moral progress is made primarily through the achievement of pre-shared goals.

---

[3] Elizabeth Anscombe, "Modern Moral Philosophy," *Philosophy* 33:124 (1958), pp. 1-19, and Alisdair MacIntyre, *After Virtue* (London: Gerald Duckworth & Co., 1981) are the two most important works of virtue ethics since Aristotle's *Nichomachean Ethics*, in *A New Aristotle Reader*, ed. J.L. Ackrill (Princeton, N.J.: Princeton University Press, 1987), pp. 363-478. Anscombe was instrumental in re-founding the discipline, though MacIntyre's formulation of it is more well-known. Other recent examples that have built on the insights of these two are Owen Flanagan, *Varieties of Moral Personality* (Cambridge, MA: Harvard University Press, 1991); Mark Johnson, *Moral Imagination* (Chicago: University of Chicago Press, 1993); and Paul Churchland, "Toward a Cognitive Neurobiology of the Moral Virtues," *Topoi* 17:2 (1998), pp. 83-96, which I will discuss in detail later.

[4] Stephen Darwall, "Introduction," in *Virtue Ethics*, ed. S. Darwall (Malden, MA: Blackwell, 2003), p. 1.

[5] Some, like Anscombe in "Modern Moral Philosophy," do seem to want to reject traditional normative and metaethical questions about moral rules and their grounding; however, others like MacIntyre in *After Virtue* and arguably Churchland do not. Aristotle in the *Nichomachean Ethics* certainly does not.

Such goals as reducing violence between individuals can be considered moral even if the full meaning of morality hasn't quite been resolved.

Thus, modern virtue ethics falls in a strange place; it is the mutt of moral theory, with ancestors in both the theoretical and the practical traditions. But it is the approach that – with the help of sciences like biology and psychology – has the most potential for growth. That's because the "what" and the "why" of moral theory – normative ethics and metaethics, respectively – are good questions indeed, but hard ones, perhaps even impossible ones; on the other hand, the "how" has been neglected for centuries, and with the advent of rigorous human science, can be answered in new and exciting ways. In any case, the "how" has some catching up to do.

Biologicizing Ethics: A Brief History

"Old beliefs die hard, even when demonstrably false." - E. O. Wilson

Unfortunately, the influence of the biological sciences on moral theory has not always been a positive one. When the theory of evolution by natural selection was just starting to become popular, it was almost immediately taken up by advocates of a so-called "social Darwinism," who used it to legitimize bigotry. Inferior races didn't need to be tolerated, for social life was simply the survival of the fittest, and the strong were *supposed* to defeat the weak. The repugnance of such a philosophy is derived from the naturalistic fallacy – the belief that whatever is natural must also be good. This is an unwarranted inference; the fact that natural selection operates on species does not mean that such selection is a good thing. Bridging the is/ought divide is not something to be taken lightly, and Social Darwinists didn't even try.

Though these bigots were eventually rebuked, the lure of the naturalistic fallacy did not fade. Since the philosophy of social Darwinism fell out of favor, one of the naturalistic fallacy's most prominent advocates came with E. O. Wilson's *Sociobiology*, which burst onto the scene in 1975 with its bold attempts to reduce social phenomena to their evolutionary, biological roots. One of these projects was "biologicizing" ethics. "The time has come," he wrote, "for ethics to be removed from the hands of the philosophers and biologicized."[6] He thought that because ethical standards rely on emotive brain regions, they could not possibly be objective.[7]

Philosophers, however, were skeptical of the man who claimed to have solved the hard problems of morality without engaging their traditions. The American philosopher Philip Kitcher was one of Wilson's major critics in the 1980s and 1990s, and he outlines the four paths he sees open for sociobiology in a paper called *Four Ways*

[6] Edward O. Wilson, *Sociobiology* (Cambridge, MA: Harvard University Press, 1975), p. 562.
[7] Wilson, *Sociobiology,* p. 564: "No single set of moral standards can be applied to all human populations."

of "Biologicizing" Ethics.[8] Though his direct target is sociobiology, his paper is an excellent example of the critiques that must be answered by anyone who attempts to bring science into ethics. In short, Kitcher attempts to show that without committing the naturalistic fallacy, there is very little that biology can add to ethics, and that most energy spent in such efforts is wasted.

The first task that Kitcher sees biology as potentially capable of is "explaining how people have come to acquire ethical concepts;"[9] that is, mining the historical origins of morality in the human species. Kitcher thinks that this is a legitimate role for biology to assume, but he is doubtful that it will be more important in this project than social or cultural history. Additionally, he scoffs at the notion that a genealogy of morals, even one tracing it back to our genetics, could by itself diminish the relevance or objectivity of ethical statements. There have been many other genealogies, he says, and many of them more convincing than the evolutionary one; while these are indeed challenges to the objectivity of morals, they are challenges that were acknowledged and at least partially answered long ago.[10] So while this first way of biologicizing ethics may be legitimate, it is far from providing new insights and will not bear ripe analytical fruit.

The next path Kitcher clears for sociobiology is the prospect for science to "teach us facts about human beings that, in conjunction with moral principles we already accept, can be used to derive normative principles that we had not yet appreciated."[11] For example, the science of happiness might alter the way in which a Utilitarian weighs outcomes. Kitcher accepts this too as a legitimate enterprise but also tries to show that using the insights of human sciences in this way is neither novel nor particularly productive. Other disciplines – environmental science and economics, for example – have long been used for this same goal, and may indeed be better at it. Wilson's "insights," and those of many evolutionary biologists, rarely present genuinely new information about human nature; they typically only provide scientific explanations of phenomena we have acknowledged for millennia.[12] Men, for example, are on average less committed to relationships than women because of the evolutionary incentives operating on each sex – but, Kitcher says, so what? Anyone who listens to standup

[8] Philip Kitcher, "Four Ways of 'Biologicizing' Ethics," in *Evolution und Ethik*, ed. K. Bayertz (Stuttgart: Reclam, 1992), pp. 439-450.
[9] Kitcher, "Four Ways of 'Biologicizing' Ethics," p. 440.
[10] For example, Nietzsche's "Genealogy of Morals" is a historical genealogy that many find convincing as a rebuke of Christian morality. Alternatively, Hume's theory of moral sentiments attempts to dissect the moral decisions that we make and attributes their source to the passions, not to an objective, rational standard. Some reject the truth of such genealogies entirely, but others see multiple layers of causation as an explanation for how they could be true and still fail to invalidate moral objectivity.
[11] Kitcher, "Four Ways of 'Biologicizing' Ethics," p. 440.
[12] Ibid, p. 442.

comedy knows that this is hardly a revolutionary finding. Thus, sociobiology here too is largely useless.

Kitcher's "third way" is to see biology as "the key to metaethics" – that is, using it either to ground the objectivity or to prove the subjectivity of moral statements.[13] If one is trying to prove a subjective metaethics, then Kitcher's response to the "first way" – that a genealogy of morals cannot by itself diminish the objectivity of ethical statements – applies here as well. If one is attempting to ground ethics in biological truths, then Kitcher is well-prepared to cry foul. He compares the acquisition of moral truths to learning mathematics: our minds evolved to be capable of doing math, after all, but that doesn't mean the truths of mathematics are dependent on our existence or grounded in our minds. Morality could be similarly objective, even if our minds did evolve towards appreciation of its truths. Surely, concedes Kitcher, there are hard problems in metaethics, but sociobiology doesn't provide any new arguments.[14]

The final direction in which Kitcher sees sociobiologists headed is towards "teaching us new fundamental principles"[15] of morality. Kitcher sees this possibility as fundamentally absurd. Wilson seems to have no qualms about committing the naturalistic fallacy: in attempting to reach new moral principles, he seeks to derive the command that we should do whatever it is that will further the *homo sapiens* gene pool, but gives no evidence other than the nature of gene-selection that this is a legitimate moral principle. Kitcher sees this as irresponsible and unworthy of philosophical recognition.[16]

Thus, Kitcher has laid forth his four ways of biologicizing ethics and deemed two of them acceptable yet insignificant, and two of them entirely incoherent. I have a few minor objections, [17] but his basic point still stands: sociobiologists bring up some age-old dilemmas in a different light, and certainly this could be a valuable contribution, but they overreach in a number of ways that discredit the rest of their work. First, they make some downright foolish claims that wouldn't pass muster in philosophical circles for a moment. And second, their legitimate claims aren't fundamentally groundbreaking in the way they are often framed. Is Kitcher right?

---

[13] Ibid, p. 440.
[14] Ibid, p. 443.
[15] Ibid, p. 440.
[16] Kitcher, "Four Ways of "Biologicizing" Ethics," p. 444.
[17] For example, the comparison of moral knowledge to mathematical knowledge is tenuous, because it is not the mathematical truths *themselves* we are genetically predisposed to believing, but rather the mechanisms by which we may investigate them. We are born not only with the *capacity* for moral argument, but also with certain moral *beliefs* already on their way to expression. An evolutionary predisposition to believe that killing is wrong is not comparable to our disposition to count and add objects; rather, it is more like a disposition to believe that 2+2=5.

The answer is both yes and no. First, yes: almost everything Kitcher says about sociobiology's flaws is correct. Wilson's focus, as a scientist, is on discovering facts with predictive value – he asks, for example, how the judgments we commonly refer to as moral are actually made, and why. Studies that focus on this are interesting, but Kitcher is right that they say absolutely nothing to the Kantian, who thinks that the only pure moral behaviors are those caused solely by an act of a free will. The word "moral" in these two cases means entirely different things.

But Kitcher is also wrong in one important way. The predictive and the prescriptive approaches to the moral are two sides of the same coin, and the hostility that they have shown each other belies missed opportunities for cooperation. Though elusive, these opportunities lie in the common goal that all parties share – greater moral righteousness in society. Kitcher has neglected a fifth way of biologicizing ethics: using the insights of biology to aid us in the accomplishment of our shared objectives. While this is not part of the scope of contemporary philosophy, it has roots in the ancient approach to morality, and it could be where we make the most progress in moral thought over the next century. Discoveries about moral learning might radically alter the way we teach morals; behavioral studies about social moral cognition could fundamentally change our expectations of social and political systems.

There are two major critiques of this position. First, this is very similar to what the ancient "practical" ethicists did, and the modern "theoretical" response appears to be quite effective at dismantling the practical approach: how can we change moral behavior for the better if we don't know what really is better? And second, it appears that Kitcher has already recognized in his "second way" the possibility that we might elaborate moral principles we already share, given biological insights into human nature, and dismissed this possibility as irrelevant. His objection is not that such principles are invalid, but that all attempts at this so far have been relatively fruitless.

Both are convincing critiques. But there is one assumption about our moral landscape present in both that I must object to – that is, the presumption that if we don't know everything about morality, then we must know nothing. The major moral tragedies of this world are not ambiguous. It doesn't take a moral genius to know that the Holocaust was something we should probably avoid in the future, and no one seriously advocates for the murder and rape of innocent civilians. The moral work to be done here is not in defining our moral goals, but in achieving them – whether or not we agree on the justification of those goals, we share them nonetheless. We don't need to postulate a definitive moral system in order to learn something from biology about how to achieve such basic moral goals.

I am sympathetic to philosophers who are very protective of their discipline, after all of the misguided attempts to biologicize ethics: scientists should concede that their genealogies don't magically resolve debates about the objectivity or content of morality. But philosophers should also admit that there is moral progress to be made with the help of science, even while we still disagree about the objective truth of the

moral goals we share. It doesn't matter whether or not such things are provable; we can still unlock the gates of moral progress – however each of us defines it. Virtue ethics is the key we have held unwittingly in our hands for over two thousand years.

Churchland's Virtue Ethics and the Promise of a Solution

"Nothing is more practical than a good theory." – Jeffrey Alexander

In the overall project of bringing the science of morality to bear on moral philosophy, the American philosopher Paul Churchland's 1998 argument in *Toward a Cognitive Neurobiology of the Moral Virtues* is both insightful and frustrating. It is insightful because he does the work of connecting the practical and theoretical approaches through the use of virtue ethics; it is frustrating, nonetheless, because he gets bogged down in debates over metaethics and is consequently unable to be explicit about his moral vision.[18] That said, I think there is something to be salvaged.

Churchland creates a concept of moral knowledge, which – in an odd parallel to Kitcher – he sees as equivalent to other types of knowledge, such as mathematical or musical knowledge.[19] His reason for this is that the neural processes guiding the development of "moral" concepts and the performance of "ethical" acts appear to be similar to those for other types of knowledge. They activate "neural networks," which control the organization of stimuli into learned patterns of behavior. Similar to the way faces are recognized by layers of neurons each doing specific jobs, Churchland maintains that we have a faculty of "moral discrimination,"[20] which is instantiated in a neural network similar to the visual one. In the same way that we might learn to recognize a certain arrangement of lines as an equilateral triangle or a particular pattern of paint on canvas as well-drawn perspective, we learn to recognize certain actions as "morally neutral," "morally praiseworthy," or "morally bad." According to his analysis, there are sets of moral "prototypes" that are learned – rather like the prototypical

[18] Churchland, "Towards a Cognitive Neurobiology of the Moral Virtues," p. 83. Churchland begins on an unpromising note; he conflates metaethics with the biological and psychological genealogy of morals. Metaethics, the study of whether and how moral statements can be grounded objectively, is not easily resolved by the facts of science, and claims that approach this grandeur are fallacious. Whether or not we can infer morals from facts – that is, whether or not there is a way of overcoming the naturalistic fallacy – is a hotly contested topic that deserves much more than a simple dismissal. The burden of proof is on those who would bridge the is/ought divide to show that this is legitimate, not on those who would challenge them. However, Churchland thankfully chooses to focus the substance of his paper not on barking up this fruitless tree, but instead on following in the footsteps of Aristotelian virtue ethics.

[19] Ibid, p. 84.

[20] Ibid, p. 86.

"triangle" in mathematical knowledge – through repeated examples. We are taught by moral teachers to recognize moral and immoral actions as such, receiving reinforcement for correct recognition.[21] By similar processes, we then use these appraisals to inform a parallel set of "skills" for acting upon these judgments. We learn, for example, the "structure of social space," and certain ways to "navigate it effectively,"[22] again from our role models and moral teachers. They reward us for moral responses, and punish us for immoral ones. Just as some neural networks are trained to recognize moral prototypes, other networks are trained to respond with moral behaviors.

Moral knowledge, thus conceived, is importantly neutral on the question of moral realism – that is, the view that at least some moral claims are true with respect to an objective standard. It might be that moral knowledge is comparable to mathematics, where our judgments may be deemed either "correct" or "incorrect." If you recognized an x^2 curve as an x^3 curve, you would be wrong, and your prototypes would need replacement. In the same way, judging an action as morally praiseworthy could also be objectively correct or incorrect. On the other hand, it is also possible that moral knowledge is more comparable to artistic skill and preference, where our recognitional and habitual abilities may only be judged subjectively. If you were to listen to Beethoven and Rimsky-Korsakov and pronounce them creative equals, I would likely think twice before accepting your musical recommendations, but I would recognize that my judgment can only be subjective. Judging an action as morally praiseworthy might, conversely, be more like this: a matter of taste. Churchland's description of a neural-network basis for morality does not force us to take a position on the objectivity of morality.

Churchland's next step is to describe the processes of moral disagreement, both within a mind and between minds. In both cases, there is a conflict about the alignment of an observed sequence of events with one moral prototype or another. Within an individual, for example, there may be a conflict between "protection" of a friend's feelings and "honesty" in telling her the truth. The same situation may activate two different moral prototypes, and it may not be clear to the individual which one should win out. As for argument between individuals, Churchland says:

> It is a matter of nudging your interlocutor's current neuronal activation-point *out* of the attractor-category that has captured it, and *into* a distinct attractor-category. It is a matter of trying to change the probability, or the robustness, or the proximity to a shared neural prototype-pattern of your opponent's neural behavior.[23]

[21] Churchland, "Towards a Cognitive Neurobiology of the Moral Virtues," p. 87.
[22] Ibid, p. 87.
[23] Churchland, "Towards a Cognitive Neurobiology of the Moral Virtues," p. 88.

Here, the mathematical and musical analogies come in handy again. If a student is unfamiliar with algebra, he may be conflicted about whether a curve he sees is an x^2 or an x^3; it could be that moral confusion is like this. In a mathematical argument, I convince my student that the pattern he sees is, in fact, an x^2 curve, just as I might convince a moral learner that abortion is a case of murder, not body modification. However, it could also be that the better comparison is to taste – that moral conflict about a situation is simply a matter of choosing which prototype you would prefer to attach it to, with no right or wrong answers. Moral argument would be akin, on this conception, to convincing someone that the Beatles are better than the Stones.

Churchland stumbles when he attempts to graft these skills onto so-called "virtues." So far, we have only had to maintain that moral knowledge and moral argument are those neural processes that relate to moral issues – no particular pattern or behavior has been identified as the correct one. But when we start to label certain skills as "virtues" and others "pathologies," as Churchland does, this is a foray into the complexities of traditional moral theory. He explains his use of these terms as an "inextricable mix of *functional* understanding and genuine *moral* understanding as brought to bear by common sense and the civil and criminal law,"[24] and at another point, he cites the "difficulty" – but not impossibility – "in clearly specifying any [moral character] as being uniquely ideal."[25] This ambiguity about the status of moral claims is frustrating, especially because it is exactly this sort of discussion that prevents Churchland from effectively challenging dominant conceptions of morality. Instead of dancing around the question of moral realism, he should have stated confidently that you don't *need* a definitive stance on the complex issues of normative ethics or metaethics in order to make progress. Indeed, the beautiful insight of the practical approach is that we needn't resolve every theoretical dilemma in order to begin unraveling the difficulties of human existence.

If, as human beings, we could calmly and rationally choose a set of moral rules to follow, and keep those rules for the rest of our lives, life would be good, and the only work to be done would indeed rest with the finer points of moral theory. But that's not our nature as moral creatures. Instead, we constantly violate even the most basic rules we set for ourselves. The major challenge to moral behavior on a daily basis isn't determining the correct course of action, but actually *doing* what we say we want to do. As a result, Churchland notes, "moral character is not something – it is not *remotely* something – that can be acquired in a day by an Act of Will or by a single Major Insight."[26] Rather, it is *in its very nature* something that must be learned over many years of repetition and habituation. Because of the way humans come to learn the thought processes, emotional reactions, and behavior patterns associated with morality, virtue

[24] Churchland, "Towards a Cognitive Neurobiology of the Moral Virtues," p. 89.
[25]Ibid, p. 89.
[26] Ibid, p. 89.

ethics is a better strategy for making moral progress than either deontology or consequentialism. Those systems focus almost exclusively on well-developed rational responses to moral stimuli, while virtue ethics strives to build well-trained rational, emotional, and behavioral reactions for a holistic moral response that is more thorough and more reliable than an entirely rational approach.

Given the way humans actually come to moral decisions, we cannot simply memorize a maxim and expect ourselves to apply it in every situation. Instead, the nature of moral action and motivation *must* be taken together with moral thought to form any holistic conception of morality. This is the single most important contribution of virtue ethics; this is what is left behind by deontology or consequentialism. With his neurobiological exploration of moral decision-making from beginning to end, then, Churchland has chipped the first crack in the hitherto unyielding wall between science and ethics.

Refining the Practical Approach

"We can't fix what we don't understand." – Paul Churchland

Some questions remain, however, for any eager advocate of such bridge-building. What sorts of substantive changes might we actually make if we accept the practical approach? Many answers lie in the disciplines of education, law, and psychology, which brings up another challenge: why should philosophers care about the practical approach? Even if it is valuable to pursue, might it not be done more efficiently without the slow and delicate process of philosophical scrutiny? I'll answer both challenges in turn.

The most obvious substantive change to be made lies with our approach to moral education. We should train moral learners – be they children or adults – *habitually*, not just intellectually.[27] Even if they have what we consider to be perfect moral prototypes, they may not recognize where to apply them, or they may lack the motivation and knowledge to do so. Of course, there have been different approaches to moral education in the history of philosophy, and some have resembled the kind Churchland would advocate. But neuroscience adds new force to their arguments and shows us that moral education cannot be an afterthought; any good moral theory must place the processes of moral learning at the center of its consciousness. As it is, parents and teachers assume that they have done their duty simply by telling their children what the rules are; many religious and spiritual groups focus on conversion experiences or moral awakenings that will permanently change behavior. Unfortunately, research on moral learning says this probably isn't enough: those hoping to help moral learners must be honest with their students about the difficulties and challenges involved with building

[27] Churchland, "Towards a Cognitive Neurobiology of the Moral Virtues," p. 91.

a moral character, and be mindful that it will take more than one insight to become a good person.[28]

In the discipline of law, Joshua Greene and Jonathan Cohen would add that prophylactic policies aimed at shaping behavior through scientific understanding of its roots are both smarter and more humane than retributive, punishment-based policies. In their article, "About the law, neuroscience changes nothing and everything," Greene and Cohen argue that forthcoming research will necessitate no changes in the letter of the laws, but it will necessitate dramatic changes in their application and spirit. The major recommendation of Greene and Cohen echoes Churchland's: that we should undertake a shift "away from punishment aimed at retribution in favor of a more progressive, consequentialist approach to the criminal law."[29] Greene and Cohen recognize that this is not a new theory of punishment, but they assert that neuroscience gives this particular theory powerful new arguments. If we discover a tumor in someone's brain that caused him to become a child molester, then perhaps we will be less willing to blame him in some metaphysical sense for the crime. We still won't want him living near our children, of course, and so "punishment" of some sort is still in order, but the motivation behind our decision will be based on the consequences of that punishment rather than a desire to exact retribution for the commission of evil acts.

In psychology, there are scores of books that attempt to show how we might structure our existence in an attempt to live in harmony with our erratic tendencies instead of locking ourselves in perpetual battle against them. Gary Marcus' *Kluge* is one such book, based on the premise that the human brain, as the product of a far-from-perfect evolutionary meandering, is a fundamentally irrational and incompetent organ. We should acknowledge our ineptitudes rather than glossing over them, he says, likening his view to the mantra of Alcoholics Anonymous: "recognition is the first step."[30] He recommends thirteen structural changes we can make – ways to work around our failures. Recommendation number five, for example, is to "anticipate your own impulsivity and pre-commit."[31] We can avoid the vices of the moment simply by preventing ourselves from having the option of doing wrong in the first place.

---

[28] It should be noted here that the terms "good," "right," "moral," etc. are all used with the full knowledge that they are open to interpretation: some might view them as akin to mathematical terms, referring to objectively correct and incorrect actions, while others will view them as subjective judgments. It doesn't matter, for the purpose of this argument, how the reader chooses to interpret them.

[29] Joshua Greene and Jonathan Cohen, "For the Law, Neuroscience Changes Nothing and Everything." *The Royal Society* 359 (2004), p. 1775.

[30] Gary Marcus, *Kluge: The Haphazard Construction of the Human Mind* (London: Faber and Faber, 2008), p. 164.

[31] Marcus, *Kluge*, p. 167.

So here we have concrete recommendations for education, for the law, and for individuals in their daily lives, and we didn't need a philosopher for any of them. Why, again, am I recommending that philosophers concern themselves with fundamentally scientific pursuits? I suppose the answer depends on our conception of the purview of moral philosophy. If it is simply a theoretical discipline, concerned only with extreme hypotheticals, then it is perfectly acceptable for theorists to continue their squabbles, ignoring the advances of evolutionary biology and neuroscience. For neuroscience has nothing to say about whether or not to keep a child in perpetual agony if it makes the rest of society happy, or whether or not to lie to a Nazi. If a philosopher is willing to say that determining answers to these questions is where moral philosophy ends, then I can do nothing but register my grave disappointment.

However, if philosophers think that a coherent moral theory should go beyond recommending the right course of action in thought experiments and reach towards a fuller conception of our humanity, then the science of moral learning is unquestionably within their scope. What is needed is a revival and refinement of the practical approach to morality, instantiated in a holistic virtue ethics that resembles Churchland's. Except instead of merely de-emphasizing the questions of normative ethics and meta-ethics, this new practical ethics should explicitly set them aside, taking on instead only those ethical tasks that are uncontroversial. These tasks, after all, are quite a lot of work in themselves. Nobody would object to attempts at training kindness, understanding, independence or courage in our children – in fact, we already try to do this in our public schools. But we do it badly, and moral philosophy should share the burden of the other disciplines in changing this. Philosophers have much to contribute: if nothing else, they should lend to these scientific recommendations the coherence and force of an ethical doctrine.

Conclusion

"The philosophers have only interpreted the world, in various ways;
the point, however, is to change it." – Karl Marx

When he outlined four misguided ways of biologicizing ethics, Philip Kitcher hit nearly everything on the mark. Almost without exception, the attempts of biologists and psychologists to bring science into the ethical realm have fundamentally misunderstood the project of moral reasoning, as conceived by the philosophers. But the problem is not entirely the fault of the scientists. Nor does it really lie with the philosophers. Rather, a deep divide between the practical and theoretical approaches to morality prevents the scientists' predictive claims from having salience within the normative and meta-ethical traditions – so the insights of scientists have been largely ignored by moral philosophy. And for the sake of our place in the world, this is a shame.

Kitcher, too, failed to see across this moral chasm. Like almost all of the philosophers in his tradition, he failed to recognize that contributions to normative ethics and meta-ethics are not the only valuable contributions to moral theory. No matter what system of rules we accept, or whether we think it is grounded in objective truth, we do have common moral goals, and moral philosophers have a role in seeing those goals accomplished. Paul Churchland recognized that, but he was unable to convey it without getting trapped in the quicksand of unproductive metaethical examination. For moral progress to be made, we must bridge the chasm between the practical and the theoretical by putting aside our areas of disagreement, and embracing what science has to say about our common purpose. Ethics can be biologicized, but not in the way that Kitcher imagined. Instead of imposing its own norms, science must let ethics guide its inquiry. Moral philosophers will ask questions about how to achieve shared ends, and scientists will provide answers. But this productive dialogue can only be achieved if philosophers accept the legitimacy of scientific conclusions about moral behavior. Only then will scientists and philosophers stop talking past each other, and only then will we begin to make moral progress.

Works Cited

Anscombe, G.E.M. "Modern Moral Philosophy." *Philosophy* 33:124 (1958), 1-19.

Aristotle. *Nichomachean Ethics*. In *A New Aristotle Reader*, edited by J. L. Ackrill, 363-478. Princeton, NJ: Princeton University Press, 1987.

Churchland, Paul. "Toward a Cognitive Neurobiology of the Moral Virtues." *Topoi* 17:2 (1998), 83-96.

----- "Precis of The Engine of Reason, The Seat of the Soul: A Philosophical Journey into the Brain." *Philosophy and Phenomenological Research* 58:4 (1998), 4859-4863.

----- "Replies to Critics." *Philosophy and Phenomenological Research* 58:4 (1998), 893-904.

Darwall, Stephen. "Introduction." In *Virtue Ethics*, edited by Stephen Darwall, 1-4. Malden, MA: Blackwell, 2003.

Flanagan, Owen. "The Moral Network." In *The Churchlands and their Critics*, edited by R. N. McCauley, 192-215. Cambridge, MA: Blackwell, 1996.
----- *Varieties of Moral Personality.* Cambridge, MA: Harvard University Press, 1991.

Greene, Joshua, and Cohen, Jonathan. "For the Law, Neuroscience Changes Nothing and Everything." *The Royal Society* 359 (2004), 1775-1786.

Howard, Ronald, and Korver, Clinton. *Ethics for the Real World.* Cambridge, MA: Harvard Business Press, 2008.

Hursthouse, Rosalind (2007). *Virtue Ethics.* Retrieved December 20, 2008, from Stanford Encyclopedia of Philosophy: <http://plato.stanford.edu/entries/ethics-virtue>

Johnson, Mark. *Moral Imagination.* Chicago: University of Chicago Press, 1993.

Kitcher, Philip. "Confessions of a Curmudgeon." *The Behavioral and Brain Sciences* 10 (1987), 61-100.
----- "Four Ways of 'Biologicizing' Ethics." In *Evolution und Ethik*, edited by K. Bayertz, 439-450. Stuttgart: Reclam, 1992.

Kohlberg, Lawrence. "Stage and Sequence: The cognitive-developmental approach to socialization." *Handbook of socialization theory and research*, edited by D. A. Goslin, 347-480. Chicago: Rand McNally, 1969.

MacIntyre, Alisdair. *After Virtue.* London: Gerald Duckworth & Co., 1981.
----- "What Can Moral Philosophers Learn from the Study of the Brain?" *Philosophy and Phenomenological Research* 58:4 (1998), 865-869.

Marcus, Gary. *Kluge: The Haphazard Construction of the Human Mind.* London: Faber and Faber, 2008.

Scott, John F. *Internalization of Norms: A Sociological Theory of Moral Committment.* Edgewood Cliffs, NJ: Prentice Hall, 1971.

Wilson, Edward O. *Sociobiology.* Cambridge, MA: Harvard University Press, 1975.

The Obscure One:
Understanding Unity in the Language of Heraclitus

FAHD HUSAIN
McGill University

In grappling with the obscure nature of his writings, interpreters of the pre-Socratic philosopher Heraclitus often tend towards one of two poles. Either they choose to echo the reception Heraclitus faced in antiquity, when his peers dismissed his work as a collection of absurd riddles, or they reiterate the contemporary interpretation that sees in his work a fundamental unity despite his numerous paradoxical statements. This essay will not side with either of these readings. Instead, it will simultaneously engage both polar interpretations of "absurdity" and "unity" to re-read the characteristic "obscurity" of the Heraclitean fragments as a rhetorical strategy underscoring the irreducible *ambiguity* inherent in ontological and epistemological claims. More specifically, I will investigate Heraclitus' attempt to conceptualize "Logos," the fundamental, ontological commonality of Being, the very "order" or "essence" unifying all reality. I will go on to claim that Heraclitus advances an inherently ambiguous and somewhat "paradoxical" conceptualization of Logos, conceiving of this "essence" or "order" not as a common ratio or static unity but as a fiery, continuously-becoming *unity-in-flux*. Particularly important for my argument will be the explication of the role that language plays in Heraclitus' thought, language being the medium through which reality becomes (partially) intelligible, even as the flux of this reality invariably exceeds the limited nature of linguistic signifiers. The aim of this essay, then, is to revive and conceptualize the dimension of ambiguity in Heraclitean philosophy that polarizing readings tend to dismiss or disavow, an ambiguity that permeates both the ongoing dialogue between language and reality, and the subsequent conceptualizations of reality as they take form in the realm of language. It is precisely this play between language, reality and ambiguity that remains crucial, not only for the Heraclitean quest for wisdom, but also for contemporary attempts at ontology and epistemology alike.

> "Although this Logos is eternally valid, yet men are unable to understand it – not only before hearing it, but even after they have heard it for the first time. That is to say, although all things come to pass in accordance with this Logos, men seem to be quite without any experience of it – at least if they are judged in the light of such words and deeds as I am here setting forth. *My own method is to distinguish each thing according to its nature, and to specify how it behaves*; other men, on the contrary, are as neglectful of what they do when awake as they are when asleep."
>
> - Heraclitus, *fr 1* (emphasis added)

Introduction: Being Ambiguous or Ambiguous Being?

The pre-Socratic philosophers were never lauded for the clarity of their prose, but even in such company, Heraclitus remains an exceptionally obscure figure. Innumerable readers have struggled with the ambiguities of his thought, and there are few who can confidently claim to have discovered any definitive interpretation of his philosophy. This impenetrability cannot simply be attributed to readers being far removed from Heraclitus' context, for if this were the case, then proper contextualization and historicization would shed the necessary light on his texts. Even Heraclitus' contemporaries, co-inhabiting his context in classical antiquity, were left befuddled when confronted with the work of the "riddler"; he was, they claimed, obscure, cryptic, and often contradictory – figures as great as Aristotle and Plato saw Heraclitus as holding "extreme views that led to logical incoherence," which, simply put, meant that they had no idea what Heraclitus was on about.[1]

In contrast with the reception Heraclitus faced in antiquity, wherein his philosophy was mostly perceived as an obscure mysticism delving incoherently into various paradoxes, modern interpreters tend to read the Heraclitean fragments in a very different way. Poetics and contradictions notwithstanding, Heraclitus, it is claimed today, advances a fully unified ontology, making him perhaps the first philosopher who can be credited with developing the concept of Logos, which can rudimentarily be conceived of as the "essence" common to all reality, the underlying "ratio" ordering and unifying all Being.

This essay will not side with either of these polarized readings, as each glosses over various tensions in order to fortify a particular, and particularly reductive, interpretation of Heraclitus' thought. Rather, this essay will simultaneously engage both interpretations of "absurd plurality" and "unified coherence," doing so not only to

[1] Daniel W. Graham, "Heraclitus" in The Internet Encyclopedia of Philosophy, 2005. Retrieved from <http://www.iep.utm.edu/h/heraclit.htm>

highlight the limitations of each individual reading but also to do justice to the nuanced, multi-faceted, and somewhat paradoxical conceptualization of Logos that remains at the core of Heraclitus' thinking – a unique conceptualization that can only be understood if the rhetoric used to explicate this "unity" is not dismissed as incoherence or absurdity.[2]

The aim, then, is to read Heraclitus on his own *paradoxically unified* terms, to situate him in a hybrid space between these interpretations wherein he can be recognized as a philosopher who delves into and draws upon both paradigms but ultimately belongs to neither. It is an attempt, in other words, to read Heraclitus' "obscure" language as his continual effort to conceptualize and signify *ambiguity,* a dimension that not only constitutes an irreducible, inherent component of the ontological portrait of Being as it is painted by the Heraclitean fragments, but, as we shall see, continues to have significant reverberations on contemporary attempts at ontology and epistemology alike.

As a correlate to this pervasive ambiguity, a *methodological hybridity* emerges in Heraclitus' work as his tool of choice. It is my claim that Heraclitus simultaneously makes use of two complementary approaches to adequately address the complexities of Logos. On the one hand, Heraclitus argues that the seeker of truth must employ a *phenomenological* methodology – that is to say, a methodology built upon the contemplation of perceived phenomena – to intuit the essence of Being behind the apparent plurality of the everyday. On the other hand, Heraclitus implies that this phenomeno-ontological quest for knowledge can only be sustained via a methodological supplement akin to *deconstruction*: a self-reflexive, hyper-critical approach that continually scrutinizes the medium of language through which Being comes into focus, perpetually reworking the linguistic and representational matrices through which reality is rendered (partially) intelligible and subsequently becomes knowable. This dual approach is necessary, for though phenomenology remains crucial in experiencing and intuiting the reality of Being, it is nonetheless an incomplete methodology for adequately – that is to say, critically – signifying this intuition as knowledge, given that this latter signification inevitably occurs via language. Conversely, a criticality towards language remains inadequate on its own insofar as a single-minded emphasis on the mechanics of representation often amounts to eschewing the phenomenological contemplation of that which is being represented.

The irreducible, *irresolvable* nature of ambiguity inherent to Heraclitean ontology means that the reciprocal play between phenomenology and deconstruction continues ad infinitum, giving rise to a perpetual oscillation that constitutes the core of Heraclitus' method. Indeed, it is this very oscillation that allows the Heraclitean fragments to weave

[2] Put in more general terms, *form* cannot be unproblematically separated from *content* with the content remaining unchanged. The form in which content occurs is a itself kind of content, and remains an irreducible aspect of the overall formulation. More often than not, this interplay is either simply overlooked or deliberately disavowed.

between and weave together – to *hybridize* – essence and change, unity and plurality, intelligibility and absurdity, and language and reality. Only by tarrying with this ambiguity, and attending to this oscillation, can one re-read Heraclitus' texts as a unified collection of fragments whose commonality is the perpetual attempt to use language to contemplate, comprehend and signify the ambiguities of reality – to *know* that unity-in-flux that is the fiery Logos at the heart of Being. In the spirit of re-reading Heraclitus from and for the present, it is precisely ambiguity that I strive to keep alive, stressing this ambiguity not only as a methodological necessity for the perpetual revision of human knowledge, but, more fundamentally, as an irreducible dimension of the epistemological quest undertaken to understand the very fabric of reality.

Knowing Common Truth

For Heraclitus, most people spend their lives captivated by the apparent contingency and heterogeneity of the everyday. Somnambulists asleep in their waking lives, they are like children oblivious in their innocence, like deaf people who are absent even when present (*fr 1; fr 34*).[3] Some spend life immersed in their own subjective version of truth, living in accordance with their own "private intelligence" (*fr 2*). Others get drawn into the "many," taking as their compass the 'wisdom' of the poets and the crowd (*fr 104*). Nor do the philosophically-minded fare much better in Heraclitus' eyes. Thinkers such as Pythagoras, Hesiod and Xenophanes spend their lives and their intelligence trying to learn a great many things, investigating the various forms and complexities of Being, but because they insist upon immersing and abandoning themselves to the fleeting heterogeneity of life, they end up in no better a position than the epistemologically-deficient multitudes: their attempts amount to little more than a collection of random information, a disjointed cache that can never result in true understanding *(fr 40)*. Such thinkers, Heraclitus argues, are too dependent on the tradition they inherit, too blind to its specific foundations, and, inevitably, too "learned" to shed its prescriptions; wearing such blinders, they do little more than pass down the same problems in different forms to the next generation of "intellectuals." Effectively, the many – studious and oblivious alike – simply do not understand higher truth(s): either they remain transfixed by the illusory contingency of the everyday, or they continue to parrot the norms of an inherited, and limited, epistemology.

The true form of knowledge, Heraclitus asserts, is that wherein one achieves the understanding of a *singular Logos that is common to all Being*, a Logos that can be rudimentarily conceived as the underlying "essence" common to all manifestations of Being, the "ratio" or "order" unifying all Being *(fr 2)*. Through much effort and cultivation, Heraclitus suggests, one can slowly shed the habituated blinders of the

[3] *Heraclitus: The Complete Philosophical Fragments*, trans. William Harris, Professor Emeritus at Middlebury College, Retrieved from <http://community.middlebury.edu/~harris/Philosophy/Heraclitus.html>

everyday and contemplate this essence of existence, a cultivation that is analogous to the case of children who are initially naïve and believe whatever they are told but start to think for themselves as they grow into rational adults. On the exact logistics of this education, however, Heraclitus is characteristically ambiguous. In his contradictory style, Heraclitus suggests that those who seek wisdom and true understanding should not be listening to him in the first place, but instead, should be contemplating Being itself; for such is the only path to the wisdom that "all things are one" (*fr 50*). Yet Heraclitus is not entirely without recommendations: one must maintain a fervent skepticism towards "received wisdom"; one must ground one's inquiry of the being-of-the-world *in oneself* without isolating oneself from that which is common;[4] one must wrench oneself away from the bad teachers of the many and give reason the room to develop unchecked by the normative channels of pre-established knowledge, all the while being careful not to fall prey to a merely "private intelligence."[5]

In contrast to learning by rote or by whim, Heraclitus advances the *senses* as the initial building blocks for knowledge, emphasizing above all the eyes, ears and capacity for thought: "the things of which there can be sight, hearing, and learning – these are what I especially prize" (*fr* 55). His preference for a *phenomenological* method of inquiry is obvious here. Yet, as was noted above, maintaining a distance from the many in order to engage in this phenomenology should not amount to isolating oneself from the everyday. For if Logos is common to all manifestations of Being, then even the apparently contingent phenomena of the everyday contain traces of this ontological commonality. And as the seeker of truth remains inevitably present in this everyday context, the perceived heterogeneity of his immediate environment becomes the methodological starting point for a phenomenology which seeks to intuit that which is common to all Being. It is therefore essential that lovers of wisdom do not disregard "random" phenomena, that they use their senses to "acquaint themselves with a great many particulars," all the while taking care to not allow this acquaintance to develop into an obsessive cataloguing that mistakes the trees for the forest (*fr 35*). Because the everyday remains permeated with this common (despite appearances to the contrary), unwise people often stumble onto some semblance of universal truth, but either have no idea that they have done so, or have no inkling of how to go about interpreting it

[4] For the sake of clarity, it is important to re-emphasize that Heraclitus' use of the "common" refers to that which is *ontologically* common, i.e., that which is common in the most fundamental sense. It is that which is a *commonality* in all facets and manifestations of Being, a "common" that simply and fundamentally *is*, being automatically present in all existence and experience. In no way does this "common" carry connotations of class, refer to "the masses," the "popular," and so on.

[5] Perhaps it is because of Heraclitus' fervent emphasis on the *individual* nature of the quest for knowledge that he shies away from dictating a rigid framework for education and cultivation.

correctly.[6] In other words, they simply do not have the proper intuitive and interpretive grid through which to filter perceived phenomena and distill the truth common to all things. Deceived or overwhelmed by their senses, they "do not take heed of the things they encounter, nor do they grasp them even when they have learned about them, although they think they do" (*fr 17*). Since the capacity for thought required to intuit and understand this common essence is, in theory, available to all, what ultimately distinguishes the seeker of knowledge from the many is that s/he cultivates a mode of thinking that can recognize, behind the contingency perceived by the senses, a commonality that is the key for understanding the unity of all things (*fr 113*). In the quest for true wisdom, it is the generation of this *matrix of intelligibility* that distinguishes the wise one from an "intelligent" member of the unwise many.

Fragmented Language

In repeatedly referencing that which is ontologically common, Heraclitus appears to be advancing Logos as the single ideal form to be understood by those men who "choose one thing in preference to all else," who seek the eternal glory of absolute knowledge instead of the mortal transience of brute rewards that are the lot of the "masses [who] simply glut themselves like cattle" (*fr 29*). Yet, if there is some sort of underlying commonality inherent in everyday existence, how does one render this Logos understandable, or even intelligible? In the terms of this essay, how does one build that matrix of intelligibility that leads to understanding the unity of Being, allowing one to intuit and interpret the Logos common to the same scenes of everyday life in which others see only random contingency?

For Heraclitus, phenomenology plays a crucial part as the methodological basis for answering such questions, insofar as it is through the senses that one can develop an *intuition* of a universal oneness behind the flux of everyday particularities. But a purely phenomenological approach remains limited, and somewhat perilous: if one were to rely uniquely on one's senses when faced with the constant, heterogeneous motion of Being, there would remain the danger of falling into a purely subjective "intelligence" or a trivial, "sensual" truth about the plurality of life. This, of course, is to say nothing of the danger that arises when such subjective "truths" become conflated with knowledge per se. What is necessary here is a supplement – or rather, complement – to phenomenology, one that keeps watch over the translation of phenomenological intuition into the "truth-statements" of "objective knowledge," that ceaselessly

[5] This situation also seems to suggest that there is a necessarily accidental moment of revelation for Heraclitean enlightenment, one that is seemingly beyond human control: one must continually educate oneself to cultivate the properly rational, interpretive capability so as to be equipped for the moment when one stumbles upon the unexpected, at which point one can deploy this rational capability and catch a glimpse of divine knowledge.

scrutinizes not only the tacit assumptions that anchor these statements but, more fundamentally, recognizes the mechanisms of language that make such claims possible in the first place. Such a critical complement to phenomenology must expose purported infallibilities to alternate readings and complicate simplistic, reductive and static notions with complex, nuanced and perpetually refinable conceptualizations. It is my claim that this complement is analogous to the process of *deconstruction*: a subversive practice of re-reading whose criticality towards the limited (and inevitably biased) nature of language constitutes an essential component of the quest for human knowledge, whether this quest is understood in the context of Heraclitean philosophy or in more contemporary terms. To underscore the importance of this deconstructive aspect in Heraclitus' method, one must delve into the ambiguity inherent in his fragments, an ambiguity that, as noted above, both classical and contemporary interpretations tend to suppress and avoid.

In *fr 103*, Heraclitus suggests that "[i]n the circumference of a circle, the beginning and end are common": no matter which two points are selected, these markers of "opposition" overlap. They are logically indistinguishable along the line of the circle; their commonality reflects their identity. Further, since one could, in theory, identify an indefinite (or infinite) number of oppositional/identical points on the circle with the previous argument, the commonality of a single pair of oppositions becomes indicative of a commonality that *all* possible oppositions share: a primary, omnipresent unity underlying all possible points, pervading the entire realm of this circumferential existence. But one must consider that the circle in itself is an abstract construct, an ideal form, and these points of opposition where the circle "begins" and "ends" are abstracted from this shape. Indeed, they are idealized points made possible by a common language, which, in this case, is that of mathematics. In using this circle as an example, exactly what *kind* of underlying commonality and over-arching unity is Heraclitus trying to articulate? Does this circle somehow point to a *metaphysical and ontological* "common" underlying all worldly objects, which in turn are united in Being via this commonality?[7] Or is the circle to be read as a metaphor for an abstracted, *formal* unity, where this unity is only made possible by, and rendered intelligible in, the terms of the (mathematical) language that functions as its common background?

It is this very question regarding the nature of this "common unity" that is the focus of my analysis. Traditional readings that advance Heraclitus as a thinker of unity generally ignore the linguistic dimension of his thought, interpreting the play between oppositions as Heraclitus' attempt at highlighting only an ontological common, a Logos that the seeker of truth can purportedly know in a transparent and unproblematic way. I

[6] For it to be true that the circle highlights an ontological and metaphysical commonality underlying all worldly objects, Being itself would have to function on the principles of mathematics; only then could one establish a true relation of correspondence between mathematical abstracts and the universe at large.

do not believe, however, that Heraclitus is making such a simple claim. More specifically, I argue that Heraclitus can be read as implying that language serves an irreducible mediatory role between seeker and Being, that Being only emerges as intelligible and knowable by being appropriated via linguistic mechanisms. For Heraclitus, then, any ontological portrait of Being inevitably amounts to a *linguistic representation,* a *signification* born of the refraction of reality through the prism of language.[8]

To develop this argument, I return to Heraclitus' discussion of a common unity via *fr 60*. A contemporary translation of this fragment is: "the path up and down are one and the same."[9] This particular translation focuses on the physical path and suggests that the path, as the "objective" ground of travel, is indifferent to, and takes ontological and epistemological precedence over, the direction of each travelling individual. In other words, the apparent opposition between the distance travelled by one individual in a particular direction and that travelled by another in the opposite direction is simply an illusion which conceals their true commonality – the unity of the physical path. Both directions, therefore, *are* one and the same.[10] An alternative reading would consider an

[7] To clarify: I am not suggesting that Heraclitus is a kind of Kantian arguing for external reality as an unreachable Thing-in-itself. Rather, Heraclitus, as I will go on to argue below, is implying that (i) reality is *partially* accessible via language; that (ii) all access to reality proceeds via language; and (iii) a non-linguistic, unmediated "pure" access to reality remains impossible. The difference I am stressing here is that between language conceived of as a kind of (mental/symbolic) *veil* behind which external reality remains forever hidden, and language as a *grid*, where external reality is ordered and made intelligible in relation to pre-established co-ordinates.

[8] Geoffrey Kirk, John Raven, and Malcolm Schofield, *The Presocratic Philosophers: A Critical History with a Selection of Texts* (New York: Cambridge University Press, 2006), p. 188.

[9] Michael Stokes suggests that it is unlikely that Heraclitus was arguing for anything more than the physical commonality underlying distances travelled in opposite directions, for genius though he was, Heraclitus could not think in the terms of the dichotomies of qualitative/quantitative, absolute/relative, subjective/objective. See Michael C. Stokes, One and Many in Presocratic Philosophy (Cambridge: Harvard University Press, 1971), p. 91. I find this claim problematic: if Aristotle, for example, can think of quality and quantity, though his understanding of this dichotomy may differ from its "modern" manifestation, what makes Stokes so sure that Heraclitus was unable to do so? Stokes seems to judge Heraclitus' work from a certain "modern" perspective, wherein Heraclitus is seen as failing to reach this level of thinking precisely because he does not echo the precise binary distinctions of the historically specific framework structuring contemporary philosophy. Such a reductive reading also eradicates the possibility that Heraclitus was working with analogous concepts, or perhaps he was a philosopher for whom such concepts were recognized but not given much significance.

alternative translation of the fragment: "the *way* up and the *way* down are one and the same." Here, the subjective *perspectives* "up" and "down" are equated (or, at the very least, remain equatable). Once again, the crux of the issue remains whether one considers these opposites, up and down, as (i) actually *being* one and the same, and thus highlighting a metaphysical/ontological commonality of which they are but a part, or (ii) as perspectives that are *equal (or equatable),* thereby highlighting the common language from which they emerge. The problem is rearticulated in *fr 61,* which examines a case of a simultaneous, that is to say, paradoxical, co-presence of opposites: "Sea water is at once very pure and very foul: it is drinkable and healthful for fishes, but undrinkable and deadly for men." The same body of water is simultaneously both "pure" and "foul." If we ignore the interpretation that Heraclitus is merely being absurd, the question reappears: are these opposites to be thought of as different conventions and perspectives projected onto the body of water, or is the simultaneous presence of both opposites in one common element yet another example of how all is unified under/in Being?

In *fr 111*, Heraclitus states that "it is by disease that health is pleasant, by evil that good is pleasant, by hunger satiety, by weariness rest." Once again, and quite paradoxically, there are opposites that are somehow linked despite – or rather, because of – their seemingly distinct qualities. The traditional reading would take Heraclitus to be suggesting that it is only through the phenomenological experience of one extreme that one can recognize the (pleasantness of the) other. However, the opposition can also be translated into linguistic terms: *only because these two terms share a relational tether do they remain intelligible*. That is to say, one term of the dichotomy becomes intelligible only in relation to the (absent) other: without the creation of the signifier "pain," and its

Take the example of "deconstruction" as it is used in this essay. Clearly, I am not saying that Heraclitus of 540-475 B.C. was using deconstruction exactly like Jacques Derrida did in the twentieth century; nor am I suggesting that Derridean deconstruction was necessarily an impossible concept for Heraclitus to grasp precisely because he was born a few thousand years before Derrida. The former claim eradicates all historical specificity and enacts a simple, and highly artificial, equivalence; the latter claim is overbearingly reductive, conceiving of Heraclitus only in the specific terms and terminology of 20th century deconstruction. Both claims advance the "truth" about Heraclitus, which, given the philosopher in question, is quite a proclamation. The way out of this impasse is to deny this obsession with "truth," advancing in its stead ways of reading (and re-reading) via concepts which show reasonable analogies in Heraclitus' work, concepts which can be productively used to explicate the mechanisms of his thought for the sake of enriching contemporary philosophy. "Deconstruction" is the concept/name that this essay uses for this purpose; "phenomenology" is another concept that, although addressed here to a certain extent, remains under-theorized and could be developed to generate equally productive readings of Heraclitus' philosophy.

assignment to disease, one would have no intelligible conception of health as a phenomenological state of "pleasure." It therefore becomes impossible to *understand* one experiential state intelligibly without (naming) an-other; generally speaking, *intelligible meaning only emerges via relational oppositions between terms*. Read along such lines, this fragment implies a *common, unified realm of intelligible relationality*, another name for which is *language*.[11]

A complex combination of these oppositional tensions is demonstrated by *fr 88*:

> It is one and the same thing to be living and dead, awake or asleep, young or old. The former aspect in each case becomes the latter, and the latter becomes the former, by sudden unexpected reversal.

As the second sentence of this fragment demonstrates, the emphasis here is placed on a pair of opposites that are temporally resolved into one another even as they re/generate from one another. Each changes into the other and each emerges from within the other. Put otherwise, each is the condition of possibility for the other, and each is only intelligible in terms of the other. For example, not only are the terms of the young/old dichotomy sustained as long as they remain in opposition to each other, this tether upholds the *linguistic category* of "age" which serves as a common yardstick of relational intelligibility differentiating (naming) the "old" from the "young" on both intrapersonal and interpersonal levels. The simultaneous and continuous presence of these oppositional terms points to an underlying commonality that persists despite incessant flux, an internal consistency sustaining an overall unity despite the fact that the composite terms perpetually undergo "unexpected reversals" – once again: the realm of language.

Given that the identities and boundaries of these binaries emerge when they are named in opposition to one another, that these binaries are sustained only in relational

[10] Heraclitus can be read as being very close to Derrida here. For Derrida, this relational opposition would be repeated indefinitely and infinitely: for one term becomes intelligible only in relation to another term, which, in turn, becomes intelligible only in relation to another, and so on ad infinitum. Furthermore, all terms can generate relations with an indefinite amount of terms in different contexts, which means that the intelligible meaning of these terms continually shifts, morphing to suit the context of enunciation and reiteration. This fluid network of relationality – the Derridean "text" that serves as the very condition of possibility for intelligibility – extends indefinitely and unpredictably, retaining an infinitely open-ended movement that inevitably undercuts all absolute claims to objective truth. For given the continuously morphing nature of this linguistic realm, such "immutable," that is to say, static, truths are revealed as being nothing more than a temporarily intelligible and ultimately revisable configuration of meaning in a given context.

difference with each other, and that such signifiers only become intelligible against a common backdrop or unified framework whose structural rules allow for such relational meaning to emerge, it becomes clear that language is a crucial element at play in Heraclitus' fragments, one that no interpreter or follower of his philosophy can afford to ignore.

Yet though the above re-readings underscore the importance of language and its mechanisms in Heraclitus' fragments, one can nevertheless ask: is Heraclitus to be read as simply describing the play of linguistic conventions? When, for example, Heraclitus speaks of water being both pure and foul, does he simply mean to point out the different categories mapped onto this element? What, if anything, does Heraclitus have to say about the *materiality* of water aside from its linguistic appropriation? Cannot one discern a drive in Heraclitus' fragments to "go beyond" the linguistic, as it were? Similarly, do the categories of health, rest and hunger not have a *phenomenologically lived*, undeniably "real" aspect that cannot be simply reduced to linguistic mechanisms and processes, be they signifiers of pain and pleasure, relations of oppositional intelligibility, naming, and so on? Clearly, one cannot restrict oneself to the terrain of language; there is no moment when the oppositional signifiers of language are not in dialogue with something lying beyond the linguistic realm, no moment when oppositional categories do not correspond to, or attempt to approximate, some "real" correlative substance or state.[12] My argument here is that this extra-linguistic realm points to the presence of a materiality over and against language, a "world" that lends itself to linguistic appropriation, but, in the final instance, always exceeds and destabilizes the static categories of language.

In an attempt to explain the relationship between language and world, Patricia Curd suggests that "Heraclitus likens the possession of real knowledge to the comprehension of language, and the structure of the world to the structure of language."[13] But likening the structure of A to that of B is not to conceive of them as one and the same, nor does it necessarily mean that that they can be equated; it simply means that A and B have a certain *relation of correspondence* between them, that they are, at best, somewhat analogous. Put otherwise, there remains an irreducible, and irreducibly *ambiguous*, gap between the structure of language and the "structure" of a unified Being,

[11] This perhaps is one of the differences between a deconstructive Heraclitus and the deconstruction found in the works of Derrida. For if Heraclitus is read in the way as he is here, there must a corresponding material substance for the signifier that gives the signifier its stability, a substance/object that is always being signified by – that corresponds to – that signifier, even as this substance always exceeds the limited forms and mechanisms of signification. For Derrida, the signifier is fundamentally detached from any notion of a "present" materiality; the stability of the signifier is installed retroactively, where the signifier subsists only because it is constantly reiterated in a consistent manner.

[12] Patricia K. Curd, "Knowledge and Unity in Heraclitus," *Monist* 74:4 (1991), pp. 531-550.

wherein the static signifiers of language continuously attempt to approximate the ceaseless movement of Being, but inevitably fail to do so once and for all. Yet for all their limits, linguistic signifiers must not be taken as fictions: "night" and "day" do signify some perceived, interconnected phenomena, as do the placeholders of "young" and "old." My re-reading of the Heraclitean fragments can be rearticulated in this light: I am not only arguing that the relations of opposition and unity can be situated in, and considered in the terms of, linguistic mechanisms; I am also arguing that the ongoing play of these oppositional signifiers demonstrates how language continually establishes a relation with a fundamental, material unity that it struggles to approximate, a unity that is simultaneously rendered *partially* intelligible by language but nevertheless *always-already lies in excess of the latter's reductive signifiers*.

A brief example should be instructive in this regard. Consider the case of "divine law" in the following fragment:

> Men should speak with rational mind and thereby hold strongly to that which is shared in common – as a city holds on to its law, *and even more strongly*. For all human laws are nourished by the one divine law, which prevails as far as it wishes, suffices for all things, *and yet is something more than they are* (*fr 114*; emphasis added).

Here, using the analogy between divine and human law, Heraclitus demonstrates the excessive and ambiguous nature of reality with regards to the limited nature of language. The signifier that is "divine law" can be used to make a *comparison* with the laws of the human city. But the comparison remains limited, insofar as a truly divine law would not only transcend the finitude and fallibility of human laws, but also exceed human comprehension altogether. Effectively, the laws of divinity can thus can only be (reductively) referred to via a linguistic signifier as that which lie "beyond" – that is to say, they can only be *ambiguously analogized*.

But perhaps the fundamental question still remains: given this irreducible chasm between language and the extra-linguistic, how can one attempt to signify and know the real unity called Logos, that unity-in-flux of the real?

Fire, Logos, Unity

According to Heraclitus, the task of the knowing-subject is to pursue the singular object of knowledge, to recognize that "[w]isdom is one thing. It is to know the thought by which all things are steered through all things" *(fr 41)*.[14] Here the fragment

[13] John Burnet, 1912 English translation of Heraclitus' fragments in Philoctetes, 2006. Retrieved from <http://philoctetes.free.fr/heraclite.pdf> I have chosen this alternate translation for a particular reason. Harris uses "intelligence" instead of "thought," thereby translating the second limb of the fragment as: "to know the *intelligence* by which all things are steered through all things" (emphasis added). For the purposes of the argument that follows,

can be read in the traditional way: one should educate oneself in such as way as to uncover the one insight through which all of existence can be understood, thereby grasping the singular thought which latches onto the ontological consistency permeating all Being.[15] The task for the knowing-subject would simply be to discover this single, essential, transcendental thought.[16] Even the methodology of deconstruction could be appropriated into this reductive reading: deconstruction here would function as that which "strips away" the "illusions" of language, as a methodological stepping stone towards a pure, objective, linguistically un-mediated thought able to comprehend the order of all things.

Yet, such transcendentalism is at odds with the ambiguous, contradictory and paradoxical nature of Heraclitus' work. Indeed, Heraclitus would simply suggest that clamoring after a static, eternal thought – an ideal form – is a useless search, for "seekers after gold dig up much earth and find little" (*fr* 22). Heraclitus' arguments are much more complex: as the previous section hinted, they demonstrate a recognition of *the irreducible mediation of language in the knowing of Being*. In this regard, the deconstructive reading of *fr 41* above is particularly instructive: what is emphasized is not the knowledge of *the* thought (of all Being), but the knowledge of the *way thought thinks* – that is, through the linguistic categories, signifiers and statements of pre-established "knowledge." Heraclitus's insight here is self-reflexive: he recognizes the tendency of thought to *actively impose* a matrix of intelligibility upon the flux of the real in order to simplify the complexity of perceived phenomena, to forcibly construct and employ a framework of "knowledge" that makes sense of the chaos of the everyday. It is through such pre-reflexive procedures that apparent meaninglessness is transformed into a stable network of static objects in clear and unambiguous relation with each other, a pre-conscious "ordering" that highlights the automatic tendency of thought to establish a state in which "all things are steered through all things."

Consequently, the knowing-subject can never contemplate Logos directly, or purely. His or her understanding is always-already intertwined with the categorical and axial tendencies of thought itself. Here it should be noted that Heraclitus is not necessarily deriding the automatic tendency of thought towards order and stasis. Rather, his methodological emphasis is on highlighting this irreducible tendency of thought *and its permanent link with the categories of language,* insofar as thought automatically thinks via the linguistic categories, signifiers and statements of pre-established "knowledge."

---

the translation with "thought" is more instructive. However, one can make the same argument by recalling that Heraclitus only uses the word "intelligence" in a derogatory way, to designate and denigrate the subjective "wisdoms" of those who have turned inward and isolated themselves from that which is common.

[15] Stokes, *One and Many in Presocratic Philosophy*, p. 88.

[16] This, of course, would be the traditional reading of Heraclitus, which holds out the possibility of transparent, objective, non-mediated knowledge.

Hence the importance of deconstruction, both as that which investigates the linguistic bases of epistemological and ontological claims and as that which encourages a critical self-reflexivity. Deconstruction shifts the emphasis from a misguided attempt at "directly" contemplating a "pure" Logos and its eternal form, to a recognition of how *language inevitably comes into play in ontological "truths," how Logos is always-already pre-figured or rendered intelligible through the linguistic matrices of the knowing-subject*. It is language, as the overarching unifying space wherein these categories and statements emerge, that acts as the proxy, the matrix or the grid through which the real is rendered partially intelligible. And as that which generates the matrix of intelligibility interpreting the ontological flux of Being, the language through which Being comes into focus must be scrutinized by the seeker of absolute knowledge via a deconstructive methodology, a methodology which makes visible the relational grids of language which are unreflectively used to categorize and signify the real.

To see the full implications of this argument, consider the term "Logos." For Heraclitus, Logos is the commonality of all phenomenological experience, a "common" that remains available to all (thinking or thoughtful) participants of reality, and, as such, constitutes an object of knowledge for an "awareness which is common and shared, rather than private and subject to personal whim and fancy."[17] Logos is thus usually interpreted as the "schema" of Being itself, one which remains present in all of the latter's movements and manifestations. Yet, the term "schema" is misleading. As that which lies beyond language, undergoing ceaseless motion, Logos cannot be conceived of as a stasis, an ideal form, or a mathematical principle at the heart of (an ordered) Being. Indeed, one must apply the same self-reflexive criticality and caution developed above to the very terms of my analysis. Like all linguistic signifiers, the word "schema" is a *necessary misfit*, for in trying to catalogue, segment and extract an "object" or "essence" from the irreducible heterogeneity of reality by using a static symbol or signifier, one inevitably reduces the complexity of Being by appropriating it as a one-dimensional linguistic "object." Hence one falls into the unavoidably reductive traps of language. Yet, it is important to stress that *it is these very traps that make intelligibility possible in the first place*. Such categorical misfits and static signifiers are the very conditions of possibility for intelligibility itself, a necessary methodological (mis)step that has to be taken to initiate thought, understanding and explication. For if one were unable to render a part of reality temporarily static via a signifier, and to oppose it to another similarly extracted linguistic object, relational meaning could not emerge.[18] What must always be kept in

[14] Kurt Pritzl, "On the Way to Wisdom in Heraclitus," *Phoenix* 39:4 (1985), p. 306.

[15] A few words on this methodological necessity: one can make the same claim with some of the other terms so crucial to Heraclitus' thought. Logos, for one, is perhaps the exemplary example of a categorical misfit or a *linguistic placeholder*. What it "really" corresponds to is still up for debate, but it nevertheless forms a center or an anchor which grounds Heraclitus' thought. The "true" meaning of Logos is not clearly defined at the outset; Heraclitus' whole

mind is that any attempt to conflate this methodological necessity with Logos *per se* is nothing but an impossible attempt that seeks to reduce the irreducible ambiguity of a linguistic signifier by replacing the latter with an "objective," "transparent," and purportedly infallible "truth" – a fundamentally misguided attempt.

How, then, does Heraclitus signify Logos? As he notes, no two things – be they objects, events or persons – are exactly alike, and nothing remains exactly the same over time. By being irreducibly immersed in the flux of temporality, any given thing continuously loses itself, its "self-sameness." Indeed, when the continual flux of its material environment is taken into account, a concept such as "self-same" becomes absurd. No one steps into the (self-)same river twice; reality is an unfolding process of incessant change (*fr 91*). In order to signify this paradoxical *essence* of *change*, this Logos which sustains its internal consistency despite – or rather through – a continual and infinite modification and regeneration, Heraclitus uses the symbol of *fire*. Fire gives Heraclitus the metaphor of a *unified flux* that can adequately approximate that which is the inexhaustible source and the self-sustaining "order" of an unfolding Being, that which both provides unity to Being and is the commonality of all Being, that which simultaneously persists through continuous change while continuously becoming *as* change, and that which is therefore *both the cause and agent of change*.[19] Via the metaphor/signifier of fire, Heraclitus' conceptualizes Logos as a *continuously moving equilibrium*, going beyond the reductive notion of a static "unity" by reconceptualizing Logos as a *unity-in-flux*.[20]

Underlying the eternal cosmos, then, is the "ever-living" fire that is being kindled and extinguished in equal measure (*fr* 30). "There is exchange of all things for fire and of fire for all things" claims Heraclitus, "as there is of wares for gold and of gold for wares" (*fr* 90). A fiery Logos is the commonality that continually reconstitutes a

---

task is the search for precisely this truth, which is retroactively conferred in bits and pieces onto the signifier as his thought develops. If one were to dedicate oneself to searching for the "true" meaning of a given signifier, which is only made intelligible in relation to another signifier, which, in turn, emerges in relation to yet another signifier ad infinitum, one would be perpetually shuttled from one signifier to another. The interrelated but indefinite movement of the linguistic realm (what in Derridean terms would be called "free play") would mean that one would inevitably spiral into an infinite regression. It is precisely to prevent this sort of analytical paralysis that, as a methodological starting point, one names a certain concept, confers a certain *ambiguity* upon it, and proceeds with the task at hand, retroactively resignifying the meaning of the signifier as thought proceeds.

[16] Edward Hussey, *The Pre-Socratics* (New York: Scribner, 1973), p. 58.

[17] Here, an *analogy* can be drawn between Logos and language to *approximate* the movement of Logos in the mechanisms of language. The unifying essence of the moving equilibrium that is language is something akin to an internal consistency that persists despite undergoing vast modification. In the case of language, this internal consistency is the alphabet.

material realm of exchange, wherein Logos "exchanges" itself for/as various material things, and in the process, becomes open to linguistic appropriation. For while the fiery "essence" of Logos, the flux of Being itself, cannot be directly signified, its temporary manifestations in and as distinct material things can be marked, named and categorized. Indeed as "God is day and night, winter and summer, war and peace, surfeit and hunger; but he takes various shapes, just as fire, when it is mingled with spices, is *named* according to the savour of each" (*fr* 67)[21]; emphasis added). In other words, a particular intersection of fire and matter becomes *intelligible through language* when it is named via the mechanisms of the latter: a specific instance of Logos is (temporarily) categorized for common intelligibility where "the name pick[s] out and preserv[es] in language the determinate nature of the fire and spice mixture at the time of naming."[22] It is via a deconstructive methodology that Heraclitus' wise subject recognizes the inevitable play and irreducible role of language in these temporary significations – a recognition that intuits and understands the continuously materializing nature of a fiery Logos that the one-dimensional significations of language struggle to represent.

Beyond Everyday Limits: the Real

Heraclitus, then, does not succumb to the false dichotomy between (i) a Logos that would be completely understandable if not hindered by the meddling of language, and (ii) a language that can completely exhaust the essence of existence and simply render Logos transparent. Rather, Heraclitus implies that intelligible access to Logos is only possible via the linguistic signification of changing everyday phenomena; and that wisdom is only possible through a *continual* consideration of the dynamic interplay between the categories of intelligibility and the fiery flux of Logos that they attempt to map.

Yet, in the final instance, one remains unsure as to whether universal propensities for knowledge are actually attainable for Heraclitus. He likens the difference between a child and a man as that between a man and a god (*fr 79*). But whereas children *can* grow into rational adults, can man ever ascend to the knowledge of a god? For if "human nature has no real understanding [and] only the divine nature has it," if unqualified wisdom cannot be the lot of man, man remains forever inadequate before the workings of the fiery universe (*fr 78*). Is Heraclitus speaking here of the un-reflexive, uncritical and automatic *tendencies* of human nature, or is he, more radically, suggesting that human *nature* in itself is inevitably finite when it comes to true wisdom? Is man forever condemned to continually grasp for an ever-elusive, ambiguous reality? Is the pursuit of wisdom not revealed to be an inherently impossible task, one wherein man inevitably misunderstands and is continually misunderstood?

---

[18] Since the translation for this fragment is incomplete in Harris' collection, this fragment has been taken from the Burnet translation.

[19] Kurt Pritzl, "On the Way to Wisdom in Heraclitus," p. 312.

If one does follow this more radical reading, the important question becomes: if one avoids both the indulgent subjective aspect of "intelligence" and the biased inheritance of tradition, *to what extent* is universal knowledge possible? Amidst the ambiguity of his paradoxes and poetics, Heraclitus leaves this question unanswered. Herbert Granger puts Heraclitus' motive for the ontological pursuit in the following way: "the wisest of all may also be what pre-eminently possesses life, thereby […] indicating that wisdom and a higher form of life go hand in hand."[23] Thus humans can attain a *degree* of wisdom; they can, to some extent, awaken from the sleep of their ignorant folly, and become fully alive in the quest for achieving wisdom, thereby becoming "as much like the divine as is humanly possible."[24] For the time being, then, if not forever, the quest for wisdom proceeds via a ceaseless refinement, refining both the categories of language and the knowledge they make possible. Put otherwise, the quest for wisdom becomes a perpetual, and perpetually ambiguous, learning, an unending undertaking that significantly improves the seeker's lot.

It must also be noted that it is unclear whether Heraclitus considers himself to have reached the point of absolute knowledge. The quasi-"objective" viewpoint that Heraclitus assumes in his work, which often objectifies man and his perspective, should not be considered omniscient; at best, it is a mimicking of an ultimate objectivity that is done in order to signify – to *approximate* – a reality that exceeds the finitude of human understanding. "Wisdom is one and unique" Heraclitus suggests, "it is unwilling and yet willing to be called by the name of Zeus" (*fr 32*). As a static linguistic placeholder, the word "Zeus" cannot hope to adequately encompass and express a universal wisdom that fully understands the complexities of the unity-in-flux that is a Logos in constant becoming. A deconstructive approach here would be fruitful for contemporary philosophy insofar as it would not only argue for the kind of criticality that Heraclitus sustains towards "received wisdom" but would also promote the recognition of the limited expressivity of linguistic conventions, thereby stressing the need to rework these inherited conventions of institutionalized language so as to grapple with the fundamental lack – the *ambiguity* – of any signifier of intelligibility. In effect, this deconstructive approach would encourage a philosophical stance that involves taking language as an "object of study," wherein language is recognized as "a, if not the, primary vehicle by which perspective is established, and foundations provided, for intelligible experience of what exists."[25] This approach would constitute a perpetually critical stance recognizing that each subjective perspective inevitably has its own being-

[20] Herbert Granger, "Death's Other Kingdom: Heraclitus on the Life of the Foolish and the Wise." *Classical Philology* 95:3 (2000), p. 277.
[21] Granger, "Death's Other Kingdom: Heraclitus on the Life of the Foolish and the Wise," p. 279.
[22] Joanne B. Waugh, "Heraclitus: The Postmodern Presocratic?" *Monist* 74: 4 (1991), pp. 604-623.

in-the-world and being-in-language as epistemological blind spots, and that the gap between language and reality will always remain irreducible. And when fully transcending one's mortal, subjective, and linguistic stance in the world is recognized as impossible, the deconstructive approach would stress the necessity of engaging in the properly *ethical* task of repeatedly beginning anew, of scrutinizing the tacit assumptions and pre-suppositions of one's own language and perpetually attempting to proceed asymptotically towards the ideal of wisdom with greater and greater self-reflexivity.

It is by keeping such insights in mind when considering the paradoxes, metaphors, circularities and "illogicalities" of Heraclitus' texts that one can recognize his varied use of language as an attempt to break free from the limited logic of binary relations and the paradigms of traditional rhetoric. Effectively, Heraclitus advances the idea that the language used to approach the real must itself be highly unusual, its style shifting to accommodate its ambiguous subject matter – hence, the oscillation between relatively intelligible claims and obscure poetics.[26] It is in interpreting the "contradictory" nature of Heraclitus' texts in this way that one can begin to understand how this "obscure one" used language to go *beyond* language, as it were; how he employed rhetorical signposts to point to a Logos lying beyond phenomenal and linguistic horizons; and how he gestured incessantly towards this real unity, this unity-in-flux of the real, that is ambiguously revealed in, but always definitively exceeds, the intelligible signifiers of "knowledgeable" language.

The lord whose oracle is at Delphi neither speaks nor conceals, but gives signs.
- Heraclitus, *fr* 93

Works Cited

Burnet, John (1912). English translation of Heraclitus' fragments in *Philoctetes*, 2006. Retrieved from <http://philoctetes.free.fr/heraclite.pdf>

Curd, Patricia K. "Knowledge and Unity in Heraclitus." *Monist* 74:4 (1991), pp. 531-550.

Graham, Daniel W. "Heraclitus" in *The Internet Encyclopedia of Philosophy*, 2005. Retrieved from <http://www.iep.utm.edu/h/heraclit.htm>

Granger, Herbert. "Death's Other Kingdom: Heraclitus on the Life of the Foolish and the Wise." *Classical Philology* 95:3 (2000), pp. 260-281.

[24] Hussey, *The Pre-Socratics*, p. 34.

Harris, William. "Heraclitus: The Complete Philosophical Fragments," 1994. Retrieved from <http://community.middlebury.edu/~harris/Philosophy/Heraclitus.html>

Hussey, Edward. *The Pre-Socratics*. New York: Scribner, 1973.

Kirk, Geoffrey S., Raven, John E., and Schofield, Malcolm. *The Presocratic Philosophers: A Critical History with a Selection of Texts*. New York: Cambridge University Press, 2006.

Pritzl, Kurt. "On the Way to Wisdom in Heraclitus." *Phoenix* 39:4 (1985), pp. 303-316.

Stokes, Michael C. *One and Many in Presocratic Philosophy*. Cambridge: Harvard University Press, 1971.

Waugh, Joanne B. "Heraclitus: The Postmodern Presocratic?" *Monist* 74:4 (1991), pp. 604-6.

Between Euclid, Kant, and Lobachevsky: On the Construction of Geometrical Objects in Pure Intuition

NAL KALCHBRENNER
Stanford University

Kant's theory is compatible with non-Euclidean, hyperbolic geometry. That is, on Kant's theory, the a priori forms of space and time together with the categories of the understanding ground the possibility of hyperbolic constructions in pure intuition. To show this we first develop an interpretation of Kant's theory of geometry to the extent that it concerns the construction of geometrical objects in pure intuition. Thus we show how the a priori forms and the categories make possible Euclidean constructions in pure intuition. We then proceed to the main result. The latter is independent from some of the details of the interpretation. Under minor assumptions the result can be strengthened to the following: if Kant's theory is compatible with Euclidean geometry, it is compatible with hyperbolic geometry as well.

Introduction

Kant's theory of geometry arises from a collection of concise remarks scattered throughout his critical writings, and it is embedded into a more general doctrine of cognition. What seems clear in Kant's theory is that geometry is related *in some way* to our pure intuition of space and time and that geometrical propositions and theorems are conceived of as deriving from the a priori *construction* in pure intuition of geometrical figures. However, precise details concerning the relation between geometrical objects and pure intuition are not spelled out clearly and systematically by Kant and have become a locus of philosophical investigation.

In what follows, we inquire into the relation between geometry and pure intuition. In particular, since Euclid's geometrical objects are, for Kant, *possible* objects of pure intuition, we ask, regressively, what it is that grounds the *possibility* of geometrical constructions in pure intuition. That is, we ask *which faculties* and *to what extent* these faculties make possible, according to Kant, the construction in pure intuition of geometrical objects.

Our analysis shall show that metaphysical space and time as the a priori forms of intuition *together with* the categories of the understanding make possible and simultaneously constrain geometrical constructions. Metaphysical space and time ground the possibility of certain *intuitive features*, as we shall define them, of geometrical constructions, whereas the categories make possible certain other *conceptual features*. The intuitive features together with the conceptual features constrain a set of basic constructions in pure intuition that we shall call the three *fundamental descriptions*. On the basis of the possibility of the fundamental descriptions, we shall derive the possibility of *Euclid's* geometrical constructions, as suggested by Kant. Further, contrary to what is sometimes held about Kant's theory of geometry, we will also show that the possibility of non-Euclidean, *hyperbolic* constructions in pure intuition is derivable from the possibility of the fundamental descriptions. We shall draw some of the consequences of our analysis for Kant's general theory of geometry and intuition.

We begin, in section I, with a contextualization of the inquiry. In section II, we consider our main question concerning the possibility of Euclid's geometrical constructions in pure intuition. In section III, we take on the issue of the possibility of non-Euclidean constructions.

Section I

For Kant, geometrical propositions are *synthetic* and derived in *pure a priori* intuition. Here we restrict ourselves to a brief sketch of Kant's general argument for this claim as given in the *Critique*. In the *B-Introduction*, Kant writes:

> Just as little is any principle of pure geometry analytic. That the straight line between two points is the shortest is a synthetic proposition. For my concept of **the straight** contains nothing of quantity, but only a quality. The concept of the shortest is therefore entirely additional to it, and cannot be extracted out of the concept of the straight line by any analysis. (B16, emphasis in the original).[1]

Kant suggests that a geometrical proposition cannot be derived by mere *analysis* of the *concepts* contained in the proposition. In the quoted example, the concept of a "straight line," i.e., roughly, the general features that are *thought* of and describe a "straight line," together with the concepts of "point" and "two" do not *contain* (or entail) the further concept of "shortest line." The concept of "shortest line" is quantitative, and cannot be derived from the merely qualitative features attached to the concept of "straight line." Hence, the mere analysis of what is contained in such concepts does not suffice to derive the geometrical proposition that "the straight line between two points is the

[1] Unless noted otherwise, the following and all subsequent quotation numbering refers to the original pagination in Kant's *Critique of Pure Reason*, trans. and ed. Allen Wood and Paul Guyer (Cambridge: Cambridge University Press, 1998).

shortest"; the proposition is, therefore, not analytic. In the *General Remarks* to the *Aesthetic*, Kant continues:

> All of your effort is in vain, and you see yourself forced to take refuge in intuition, as geometry always does. You thus give yourself an object in intuition; but what kind is this, is it a pure *a priori* intuition or an empirical one? If it were the latter, then no universally valid, let alone apodictic proposition could ever come from it: for experience can never provide anything of this sort. You must therefore give your object *a priori* in intuition, and ground your synthetic proposition on this. (B64)

Since concepts do not suffice, the other source of cognition is required, i.e. intuition. By *constructing* in intuition the geometrical objects (shapes and figures) that fall under the concepts (according to the schemas), one then derives by synthesis of intuition the geometrical proposition. The proposition is therefore synthetic. In addition, Kant suggests that such derivation must occur *a priori*. Since propositions of geometry are "apodictic" or necessarily valid independently of experience, i.e. are themselves *a priori*, their derivation cannot occur in *empirical* intuition. In empirical intuition, one can derive merely *empirical* propositions that do not bear such *a priori* necessity. Hence, the derivation of geometrical propositions must occur in *a priori* intuition and, specifically, in *pure a priori* intuition.[2] This is, roughly, Kant's general argument for the synthetic character of geometrical propositions and their derivation in *pure a priori* intuition. Every step in the argument has been widely debated in the literature on Kant; in what follows, we shall concern ourselves with the *construction* of geometrical objects in pure intuition, beginning with an approach suggested by Michael Friedman.

On Friedman's approach,[3] geometrical figures and proofs are, for Kant, spatiotemporal objects constructed in pure intuition. As Friedman suggests, Kant had a conception of logic limited to monadic syllogistic logic. This limited conception further constrained Kant's view of concepts: the *intension* of a concept can, for Kant, only be *monadic*, i.e. finite. Thus, given this conception of logic and concepts and given Euclid's postulates of geometry themselves, to claim, as Kant often does, that the postulates do not imply Euclid's theorems merely by logical or conceptual analysis does not seem that

[2] In the *Transcendental Exposition*, Kant writes of intuition in geometry: "But this intuition must be encountered in us *a priori*, i.e. prior to the perception of an object, thus it must be *pure*, not empirical intuition" (B41, my emphasis). Kant does not seem to conceive of *impure* a priori intuition, or at least not in geometry. If at all conceivable, an impure a priori intuition would require the construction in a priori intuition of an object that falls under an impure concept. But, for Kant, geometrical concepts are all pure.

[3] Michael Friedman, *Kant and the Exact Sciences* (Cambridge: Harvard University Press, 1992).

unreasonable after all. In fact, monadic syllogistic logic cannot capture the essentially *polyadic theory of order* explicitly stated in modern axiomatizations of Euclidean geometry; such theory of order would imply denseness of points, allowing for a strictly logical deduction of Euclidean theorems from Euclidean axioms as given in modern formulations. But, without the possibility (only given in polyadic logics) of capturing logically or conceptually such notions of denseness, a different approach must be taken that uses only monadic logic and concepts. Further, Euclid's traditional postulates given in the *Elements* are themselves *constructive-intuitive* claims, rather than logical universal or existential claims. They capture, in fact, the *intuitive* process of *straight-edge-and-compass* constructions. Given these considerations, it is not difficult to see why (monadic) logical or conceptual analysis is not sufficient to derive the theorems from the axioms and why geometrical figures and proofs must be seen as spatiotemporal objects constructed in pure intuition. Geometrical figures and proofs are *spatial* objects, because they must be actually constructed in the spatial form of pure intuition in order to capture the *spatial* and *intuitive* features of the constructions originally postulated by Euclid. Further, geometrical figures and proofs are also *temporal* objects, because they are to be *continuously* generated by the motion of a point in time, i.e. *kinematically*; such continuous generation would allow the geometrical constructions to carry *intuitive*, non-conceptual features of denseness and, even perhaps, continuity. Finally, geometrical objects must be constructed in *pure*, as opposed to empirical, intuition, because pure intuition allows, according to Kant, for a *schematic*, idealized and "universally valid" version of such geometrical objects. Hence, pure intuition is conceived by Kant as making possible synthetic inferences from the postulates to the theorems (for which monadic logic and concepts do not suffice); that is, pure intuition plays the role of inference in a synthetic calculus of geometry. This is why, Friedman suggests, Kant insists on the spatiotemporal character of geometrical constructions in pure intuition in order to derive, not analytically merely through concepts, but *synthetically* through intuition, geometrical propositions or theorems from Euclid's postulates.

Section II

We have seen the spatiotemporal and kinematic character of the construction of geometrical objects in pure intuition. We now turn to inquire into that which makes possible, in the first place, the geometrical constructions in pure intuition and grounds their spatiotemporal and kinematic character. Specifically, we consider the following inquiries: (a) *which faculties*, according to Kant, are involved in making possible the construction of geometrical *objects* in pure intuition and (b) *to what extent* each of these faculties grounds the possibility of such constructions. Let us begin with (a). In the *A-version* of the *Aesthetic*, Kant writes:

> The apodictic certainty of all geometrical principles and the possibility of their *a priori* construction are grounded in this *a priori* necessity [of metaphysical space]. (A24)

Since geometrical figures are conceived as spatiotemporal objects constructed in pure intuition, the a priori forms of intuition, i.e. the forms of space and time, must themselves make possible such spatiotemporal constructions. That is, metaphysical space and time as the forms of sensibility are involved, to an extent yet to be specified, in grounding such possibility. Kant further writes:

> Likewise the mathematical principles do not constitute any part of this system [of the principles of pure understanding], since they are drawn only from intuition, not from the pure concept of the understanding; yet their possibility, since they are likewise synthetic *a priori* judgments, necessarily finds a place here, not in order to prove their correctness and apodictic certainty, which is not at all necessary, but only to make comprehensible and to deduce the possibility of such evident cognitions *a priori*. (B188-189)

As Kant suggests, the possibility of geometrical propositions rests on the understanding and, hence, on the pure concepts of the understanding, i.e. the categories. Since geometrical propositions are derived from geometrical constructions in pure intuition, it must also be the case that the possibility itself of the constructions likewise rests, in some way yet to be explained, on the categories. In other words, there are features of geometrical objects constructed in pure intuition that cannot be grounded in the a priori forms of sensibility, but must be grounded in the pure concepts of the understanding. It is important to note, then, that both sensibility with its a priori forms of space and time *and* the understanding with its categories are together involved in making possible the construction of geometrical objects in pure intuition. In addition, by constructing such objects, sensibility and understanding provide the latter with what Kant calls (*real*) *possibility*.[4] As Kant writes, that is *really possible*, which "agrees with the formal conditions of experience (in accordance with intuition and concepts)" (B265). Geometrical objects that are constructed in pure intuition are by definition, as it were, *really possible* objects of intuition (B268). In virtue of their being constructed in pure intuition, they agree with the formal conditions of experience given by sensibility and the understanding and make possible the derivation of *synthetic a priori* propositions.

Presumably, however, the role of the understanding here is to be distinguished from its role by which the understanding makes possible the schematic exhibition *in concreto* in intuition of the *concept* of, say, a triangle (B739). In this latter case, the concept

[4] We use the term *real possibility* only to refer to Kant's notion of *possibility* as defined and to distinguish it from our own use of the term with its common sense.

of the triangle itself first arises from a previous construction in intuition of the figure. In other words, we are not already given a priori, in some mysterious way, all the necessary concepts of geometry. We derive geometrical concepts along with geometrical propositions from the constructions of objects in pure intuition. But, the *real possibility* of the *very first* geometrical constructions, for instance, that of the basic constructions allowed by Euclid's postulates, must first be shown by way of the forms of space and time and in accordance with the only concepts already given a priori, that is, the categories of the understanding. That is, metaphysical space and time as the forms of sensibility must make possible what we shall identify as certain *intuitive features* of the very first geometrical constructions of an object, whereas the categories of the understanding must make possible, correspondingly, certain *conceptual features* of such geometrical constructions; only then will the geometrical object agree with the formal conditions of experience, in accordance with both sensibility and understanding, i.e. it will be a *really possible* object of pure intuition. Now, given that, according to Kant, Euclid's constructive postulates allow for the generation of *really possible* geometrical objects of pure intuition and given the above account of the character of such constructions, let us proceed to inquiry (b) and specify in greater detail the *extent* to which (Euclidean) geometrical constructions in pure intuition are made possible, on the one hand, by metaphysical space and time and, on the other, by the categories of the understanding.

We begin by considering *the extent* to which metaphysical space, together with the a priori form of time, are to ground the possibility of geometrical constructions. Given the kinematic interpretation of geometrical constructions as suggested by Friedman, we could roughly distinguish two ways in which such constructions are made possible by the a priori forms of space and time. *First*, since the geometrical figure constructed in pure intuition is a spatiotemporal object, the a priori forms of sensibility must, first of all, make possible the intuition of such a spatiotemporal figure. That is, metaphysical space must make possible, at least, the "original" intuition of a *point* whereas the a priori form of time must make possible the intuition of the *motion* of such point in metaphysical space.[5] Here, we do not mean that point and motion are already intuited in the a priori forms of space and time; rather, the a priori forms must ground only the possibility of such "original" intuition, in order to first make possible the kinematic construction of *any* figure at all in pure intuition. In addition, as we have seen, such motion of a point in metaphysical space guarantees that the figure is *continuously* generated. An implicit property of *continuity* is then assigned to the constructed object in metaphysical space and time, a property which, given Kant's limited conception of logic and concepts, can indeed be grasped only *intuitively*, i.e. in intuition, and not *conceptually*, i.e. through concepts. Thus, the a priori forms of sensibility must make possible, first of

[5] The strict relation, for Kant, between the a priori form of time and "the general theory of motion" is suggested at (B49). Geometrical motion which presumably underlies the more general theory of motion is related to the a priori form of time as well.

all, the *continuous generation* of a geometrical object by the *motion* of a *point* in metaphysical space (and time).

Secondly, as suggested by Parsons and Carson,[6] metaphysical space, which is "subjectively given" as *actually infinite*, grounds the *potentially infinite* iterability of geometrical constructions within metaphysical space. That is, the possibility of constructing and reiterating the construction of geometrical objects *potentially ad infinitum* – a feature that Kant individuates in Euclid's geometrical constructions – is to be grounded in the actual infinity of metaphysical space. Now, Friedman has put forward the following objection:

> The crucial question, however, concerns exactly how metaphysical space … is supposed to accomplish this grounding. Is the given infinity of space as a pure form of sensible intuition supposed to be directly seen, as it were, in a simple act of perceptual or quasi-perceptual acquaintance? Are we supposed to have direct perceptual or quasi-perceptual access to such infinity entirely independently of geometry – which access we can then use to justify or to verify the possibility of Euclidean constructions? Both of these ideas appear to be very doubtful. [7]

It seems to me that the notion of a perceptual or quasi-perceptual access to metaphysical space either to verify the actual infinity of metaphysical space or to justify, say, any other facts about the underlying geometrical structure of such space (i.e., for instance, that metaphysical space is inherently Euclidean or non-Euclidean) is indeed very doubtful. However, a different, *conceptual* approach to the actual infinity of metaphysical space seems to be suggested by Kant-Schulze:

> Thus, the geometrician, as well as the metaphysician, represents the original [metaphysical] space as [actually] infinite and, indeed, as given as such. For the representation of space (together with that of time) has a *peculiarity* found in no other *concept*, viz., that all spaces are only possible and *thinkable* as parts of one single [infinite] space, so that the representation of parts already presupposes that of the whole.[8] (Some emphasis added)

---

[6] See Michael Friedman, "Geometry, construction and intuition in Kant and his Successors," in *Between logic and intuition: Essays in honor of Charles Parsons*, ed. G. Sher and R. Tieszen (Cambridge: Cambridge University Press, 2000), pp. 187-188.

[7] Friedman, "Geometry, construction and intuition in Kant and his Successors," p. 189.

[8] Henry Allison, *The Kant–Eberhard controversy* (Baltimore: Johns Hopkins University Press, 1983), p. 176.

It is granted that metaphysical space is *intuitively* perceived as apparently *boundless*, but always as actually *finite*. Similarly, geometrical constructions within metaphysical space are intuitively perceived as having no constructive limits, but, at each stage, the geometrical constructions are intuited always as *finite*. However, metaphysical space, as opposed to the space delimited by geometrical constructions, is *thought* or *understood* to be *actually infinite*. One would immediately object to such notion by underlying the fact that, in Kant's view of a *concept*, the *intension* of a concept cannot be *infinite*. But, the passage seems to be suggesting precisely that "the representation of space (together with that of time)" is the only *concept* that "has a *peculiarity* found in no other concept." The peculiarity is identified with the fact that *all spaces are thinkable as parts of the single infinite space*. That is, *all* spaces are thought in the concept of the *infinite* metaphysical space; we would identify such concept as having *infinite intension*. Thus, it seems to be the case that metaphysical space is intuited as boundless, but conceptually thought of as actually infinite.

Given such characterization of metaphysical space, two questions are to be answered: (c) how does the concept of an actually infinite metaphysical space first arise? And (d) how is, then, the grounding of the potentially infinite iterability of geometrical constructions accomplished by metaphysical space? As a clarificatory remark that shall be of importance in the subsequent analysis, we note that, as we shall take it, the *potential infinity* of metaphysical or geometrical space and the *actual infinity* of metaphysical space can only be *conceptually thought*; what is *intuitively perceived* is only the (apparent) boundlessness in intuition of metaphysical space; of course, any *finite* portion thereof can as well be *intuitively perceived* (and *conceptually thought*). Given this remark, let us briefly address the two questions. Kant writes:

> Space is represented as an infinite **given** magnitude. (…) but no concept, as such, can be thought as if it contained an infinite set of representations **within itself** [in intension]. Nevertheless space is so thought (for all the parts of space, even to infinity, are simultaneous). Therefore [*Also*] the original representation of space is an ***a priori*** **intuition,** not a **concept**. (B40, emphasis in the original)
>
> If there were not boundlessness in the progress of intuition, no concept of relations could bring with it a principle of their infinity. (A25)
>
> Now one can only view as *infinite* a magnitude in comparison to which each specified homogeneous magnitude is equal to only a part.[9] (Emphasis in the original)

[9] Henry Allison, *The Kant–Eberhard controversy*, p. 176.

Now, we have seen that metaphysical space is *conceptually* thought as *actually* infinite. (The first passage quoted here confirms this and uses this notion, then, to argue *regressively* that metaphysical space must originally be an a priori *intuition*). However, *given as a (phenomenal) fact*, as it is on Kant's doctrine, that metaphysical space is originally the a priori form of sensibility (and not some mysteriously given concept), we want to show how the *intuition* of such metaphysical space first gives rise to the *concept* of metaphysical space. In particular, since it would be inelegant to simply postulate as "subjectively given" the *concept* of an *actually* infinite metaphysical space, we want to explain how the "boundlessness in the *progress* of intuition" of metaphysical space first gives rise to the *concept* of an *actually* infinite metaphysical space. The second and third quoted passages above will help us explain this, as follows. We are given that there are no bounds in the intuition of metaphysical space. Hence, one can intuit metaphysical space as of any *arbitrary finite* extension. Given such finite intuition, one conceptually thinks any such space as *actually finite*. In addition, given that the extension of any such space is *arbitrary*, one conceptually understands any such space to be *potentially infinite*.[10] That is, any such space can then be *conceptually thought* as *potentially infinite* and also as *actually finite*. Since it is *actually finite*, any such space is also thought as *only a part* of some *higher magnitude.*[11]Since such space is of *arbitrary* finite extension (and *any* such space is thought as only a part of some higher magnitude), such higher magnitude must be of "some *infinite* kind." Further, since *any* such space is *also* thought as *potentially infinite* (and part of some higher magnitude of an infinite kind), such higher magnitude must be thought as *actually infinite.*[12] It arises thereby the concept of an *actually infinite* metaphysical space. In other words, the boundlessness of the intuition of metaphysical space gives rise to the concept of the actual infinity thereof, as suggested by the second passage quoted above. This is how the representation of metaphysical space is "subjectively given as (actually) infinite on the side of the thinker": the concept of the actual infinity of metaphysical space is an immediate (phenomenal) consequence of the *subjectively given boundlessness* in the intuition of metaphysical space. This would presumably be the suggested reasoning underlying an answer to question (c).

Given now that we intuit metaphysical space as boundless and we think it as actually infinite, it is not too difficult to answer question (d) and explain how

[10] The potential infinity in our intuition of metaphysical space cannot simply be perceived, since only a finite number of finite intuitions of metaphysical space can be perceived. However, given the arbitrariness of the extensions of such spaces, one can conceptually think the potential infinity (but, of course, actual finiteness) of any such space perceived as finite.

[11] Henry Allison, *The Kant-Eberhard controversy*, p. 176.

[12] If such "higher magnitude" were merely thought as *potentially infinite*, then it would be thought as *actually finite*, and itself only a part of some other "higher magnitude."

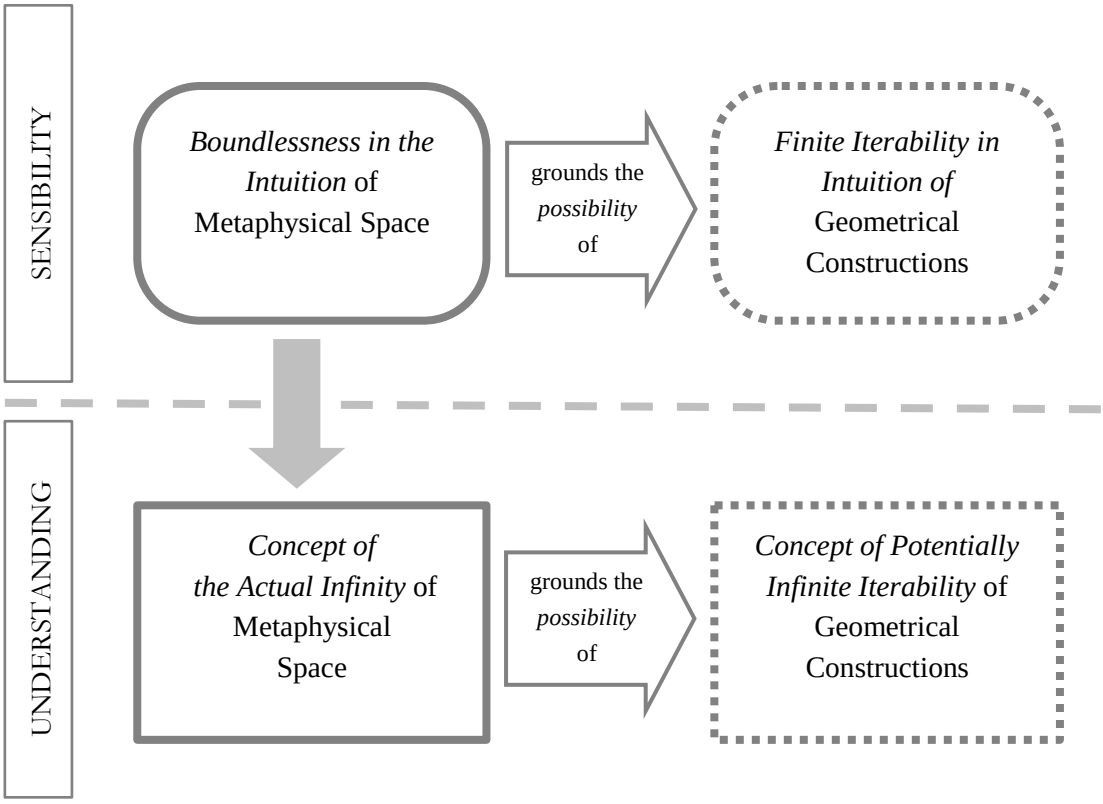
metaphysical space as the a priori form of sensibility makes possible the potentially infinite iterability (conceptually understood) of geometrical constructions. Kant suggests:

> To say, however, that a straight line can be continued infinitely means that *the space in which I describe the line is greater than any line which I might describe in it.*[13] (Emphasis in the original)

The boundlessness in the intuition of metaphysical space alone might not suffice to *ground* or *make possible* (the concept of) the potential infinity of geometrical space. In such a case, our ability of iterating geometrical constructions without apparent bounds in spatial intuition would be that which grounds the boundlessness of the intuition itself, i.e. that of metaphysical space; but this is not what Kant seems to be suggesting. On Kant's view, the intuited boundlessness of metaphysical space and, in addition, the concept of the actual infinity of space derived from such intuition ground the potentially infinite iterability of geometrical constructions. Specifically, restricting ourselves to Kant's example above, we could explain such grounding as follows: Within metaphysical space, we *intuitively* describe a line of some arbitrary finite length. We would like to show that the extension of such a line is understood to be potentially infinite. Now, the geometrical space described by the line is *intuited* as some *finite* portion of an equal or greater (but also finite) delimitation of metaphysical space. Hence, by virtue of the finite intuition, such geometrical space is also *thought* as having finite extension. Since metaphysical space is thought as actually *infinite* (in virtue of its concept), metaphysical space is then clearly thought as *greater* than the finite geometrical space delimited by the line. Further, since the length of the line was arbitrary, it follows that metaphysical space is understood to be greater than *any* line *intuitively* described in such space. Hence, as the quote suggests, it follows directly that a line is *thought* as being (or said to be) extendable potentially *ad infinitum*. In other words, the given boundlessness of the intuition of metaphysical space grounds or makes possible the intuition of a line of arbitrary *finite* length; this seems to be phenomenally uncontroversial. The *concept* of the actual infinity of space, on the other hand, grounds or makes possible the understanding that such line can be extended potentially *ad infinitum*. Indeed, as we have seen, the fact that some line as described in intuition is potentially infinite cannot be, as it were, perceptually intuited, because one can describe in metaphysical space only a finite number of lines of finite length. However, the potential infinity of a line and, likewise, the potential iterability ad infinitum of any geometrical construction in pure intuition can only be *thought* or *understood* and derived as above with the aid of the concept of the actual infinity of metaphysical space. This seems to me the way in which metaphysical space—as a subjectively given, boundless pure intuition and as an actually infinite concept—grounds and makes possible the intuitively finite (but understood to be potentially infinite)

[13] Henry Allison, *The Kant-Eberhard controversy*, p. 176.

iterability of geometrical constructions in pure intuition (see diagram). We conclude here our discussion of question (d).



Despite the role played here by the understanding, we call the continuous generation by motion of a point and the (concept of the) potentially infinite iterability the *intuitive features* of geometrical constructions that are grounded in (or derived from) the a priori spatial and temporal forms of sensibility. This is the extent to which metaphysical space (and time) make possible geometrical constructions in pure intuition. The above brief discussion shall be the first part of our answer to question (b). That the intuitive features are *necessary* features to make possible Euclid's *straight-edge-and-compass* geometry in Kant's pure intuition seems evident. However, that the intuitive features are *sufficient* to make possible only Euclid's geometry and not other kinematically constructed figures in pure intuition, as Kant would have it, is not clear at all. *A fortiori*, it seems to be the case that there must be other features to constrain the a priori construction of (exclusively) Euclid's geometrical figures in pure intuition. Let us then turn to investigate the extent to which the *understanding* and the *categories* make possible

geometrical constructions in intuition, or more specifically, how they ground the possibility of the *conceptual features*, as we shall call them, of the very first geometrical constructions. This inquiry shall concern us for the rest of this section.

For Kant, *productive imagination* represents or constructs geometrical objects in accordance with the *categories*. Here we consider that faculty, which, for Kant, is "an effect of the understanding on sensibility" and will help us explain how the categories of the understanding govern geometrical constructions in pure intuition. Kant writes:

> ***Imagination*** is the faculty for representing an object even **without its presence** in intuition… [I]nsofar as its [the imagination's] synthesis is still an exercise of spontaneity, which is determining and not, like sense, merely determinable…imagination is to this extent a faculty for determining sensibility *a priori*, and its synthesis of intuitions, **in accordance with the categories**, must be the transcendental synthesis of the **imagination**, which is an effect of the understanding on sensibility and its first application (and at the same time the ground of all others) to objects of the intuition that is possible for us… Now insofar as the imagination is spontaneity, I also occasionally call it the **productive** imagination. (B151-152, emphasis in the original)
>
> On this successive synthesis of the productive imagination, in the generation of shapes, is grounded the mathematics of extension (geometry) with its axioms. (B204)
>
> [M]otion, as **description** of a space, is a pure act of the successive synthesis of the manifold in outer intuition in general through productive imagination, and belongs not only to geometry but even to transcendental philosophy. (B155)

As Kant suggests, *imagination* is the faculty that represents or constructs some object in intuition. That is, imagination constructs in intuition the object corresponding to some specific *concept* of the object. Now, if such activity of construction occurs *a priori* in *pure* intuition (not in empirical intuition) and in accordance with *pure* concepts of the understanding (not merely empirical concepts), then such activity of a priori construction is made possible by the *transcendental synthesis of the imagination*. Thus, it is the transcendental synthesis of the imagination, also called the (synthesis of) *productive imagination*, which constructs an object in *pure* intuition and in accordance with the *pure* categories. [14] As the second passage suggests, one class of such objects constructed a

[14] We take it that Kant is really suggesting two roles of the synthesis of the productive imagination, one in the generation or construction of geometrical shapes (or generally objects) in pure intuition and the other in the derivation of some geometric cognition from such constructions in intuition. These two roles, however, presumably amount to two

priori by the productive imagination is the class of *geometrical objects*; productive imagination *generates* or, as Kant puts it elsewhere (B196), *draws* shapes in metaphysical space (and time) as the a priori forms of sensibility. In order not to be a "mere play of the imagination," such geometrical constructions must indeed occur in accordance with the pure categories of the understanding. Through this activity, then, the productive imagination "grounds geometry," i.e. it makes possible *synthetic a priori* propositions of geometry (that are derived from the constructions). As the third passage further suggests, it is rightly by the *motion* of a *point* in space, as a *description of space*, that the productive imagination constructs *kinematically* spatiotemporal objects of geometry; this is an additional confirmation of the kinematic account of Kant's geometrical constructions. Illustrating figuratively such activity of the productive imagination, Kant writes:

> I cannot represent myself a line, no matter how small it may be, without drawing it in thought, i.e. successively generating all its parts from one point, and thereby first sketching this intuition. (B203)
>
> We cannot think of a line without **drawing** it in thought, we cannot think of a circle without **describing** it. (B154, emphasis in the original)

Thus, by governing or constraining the activity of the productive imagination, the categories of the understanding make possible the kinematic constructions of spatiotemporal objects of geometry in metaphysical space and time as the a priori forms of sensibility. As we have argued above, the categories of the understanding must govern or make possible the *very first* constructions of geometrical objects. We are then given an initial explanation of how this occurs: through the activity, of course, of the productive imagination. But how exactly are we to conceive of the categories as governing such constructive activity? In particular, since it is clear from (B202 and ff.) that the categories involved in geometrical constructions are those of *quantity,* namely *unity (Einheit)*, *plurality (Vielheit)* and *totality (Allheit)*, how are we to think of these three categories of quantity as governing and making possible geometrical constructions? This is the inquiry that we take on next.

The categories of *quantity* make possible three fundamental descriptions of space by motion of a point. A specific remark by Kant shall help us explain the relation between the categories of quantity and the motion of a point in pure intuition, a remark that, despite (but, also, because of) its location in the *Metaphysical Foundations of Natural Science*, at the very end of the chapter on *Phoronomy*, "has its uses only in transcendental

aspects of one and the same synthetic activity of the productive imagination. That is, the two roles occur simultaneously and are simply two different ways of looking at the same synthetic activity.

philosophy" and explicates "the determination of all possible motion as a *quantum*." As suggested at (B204) and also by Friedman,[15] *quanta* are, for Kant, the objects of intuition as magnitudes with which *geometry* is properly concerned. This remark reads:

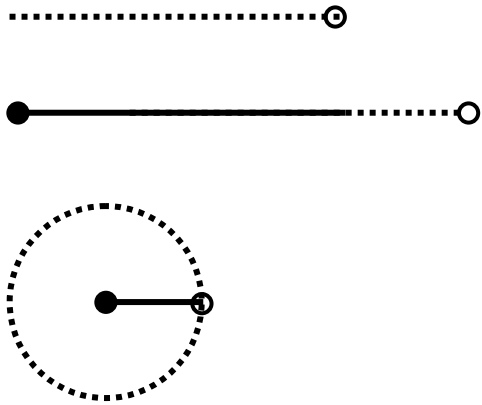
> [These are the] three moments suggested by [the structure of] space: *unity* of line and direction, *plurality* of directions in one and the same line, and the *totality* of directions, as well as lines, in accordance with which the motion may occur, which [structure] contains the determination of all possible motion as a quantum…This remark has its uses only in transcendental philosophy. (MFNS, 495)[16]

We note first that the three moments or spatial descriptions by motion of a point are suggested (*an die Hand geben*) by the structure of space. That is, space as the a priori form of sensibility itself makes possible the three spatial descriptions (i.e. the *intuitive features* of such descriptions including the possibility of *the motion of a point* itself). At the same time, however, the precise *conceptual* determination of the spatial descriptions is made possible only by the categories of quantity of the understanding. As we have seen above, it is through the faculty of the productive imagination that the understanding with its categories precisely governs such kinematic descriptions of space. Thus, drawing from the quoted passage (and generally from Kant's intended reading as suggested by the chapter on *Phoronomy*), we obtain the three fundamental geometrical descriptions of space as follows. *First,* "*unity* of line and direction": a point describes one line moving in one direction; given *unity* of direction, the line is *understood* to be straight. *Second*, "*plurality* of directions in one and the same line": a point extends a given (straight) line moving in a direction *opposite* the direction of a given point on the line; since there is a *plurality* of directions, but since the given line can only be straight (by description I), the point that extends the line can move only in the exactly *opposite* direction of some given point on the straight line. *Third*, "*totality* of directions, as well as lines": given a center-point and a segment emanating from the center-point, the segment rotates over the totality of lines that pass through the center-point or, simultaneously, a point at the other end of the segment moves in a totality of directions, describing a circle (see figure).[17]

[15] Friedman, *Kant and the Exact Sciences*, p. 107.

[16] This notation refers to the numbering in *Metaphysical Foundations of Natural Science* [MFNS], trans. and ed. Michael Friedman (Cambridge: Cambridge University Press, 2004).

[17] Circular motions are not allowed in phoronomy, because of the variable of speed (MFNS, 483). However, as *geometrical* motions, the descriptions of circles are given by a totality of directions, as Kant suggests: "A body [or point] moving in a circle changes its direction continuously, in such a way that it follows all possible directions" (MFNS, 483).



These are, for Kant, three descriptions of space that, as I take it, can be made *conceptually* precise a priori by the pure categories of quantity, namely unity, plurality and totality. That is, the productive imagination can construct *a priori* in accordance with the categories in pure intuition the above three fundamental descriptions of space, in a way that precise *conceptual features* govern their constructions, i.e. the lines are understood to be straight, the center of the circle is understood to be "equidistant" from any point on the circumference, etc.[18] The intuitive features of such descriptions, i.e. their spatiotemporal construction by motion of a point, their continuous generation and their infinite iterability, are made possible by the metaphysical space and time underlying pure intuition in which they are constructed, as we have seen above. Simultaneously, the conceptual features inherent in such descriptions are made possible and conferred on them by the categories of quantity of the understanding, which govern the constructive action of the productive imagination. Together, the a priori forms of sensibility and the categories of the understanding give *real possibility* to these three fundamental descriptions of space. That is, the objects constructed according to (repeated applications of) the three fundamental descriptions (which are obviously comparable to the process of straight-edge-and-compass construction) are *really possible* objects of pure intuition; that these objects agree with the formal conditions of experience as given by sensibility and by the understanding is here shown by the fact that their *intuitive* and *conceptual* features are made possible, respectively, by precisely the forms of sensibility and the categories of the understanding. It has been argued elsewhere by Lisa Shabel that there is, on Kant's view, a pure pre-geometric concept of space, which "warrants and constraints intuition of finite regions of space" and "provides a governing principle for all spatial construction, which is necessary for mathematical demonstration."[19] We identify such "pre-geometric concept of space" precisely with the pure categories of

[18] Given a notion of "equidistance" or "congruence" that one would consider Euclidean.
[19] Lisa Shabel, "Reflections on Kant's Concept (and Intuition) of Space," *Studies in History and Philosophy of Science* 34:1 (2003), pp. 45-47.

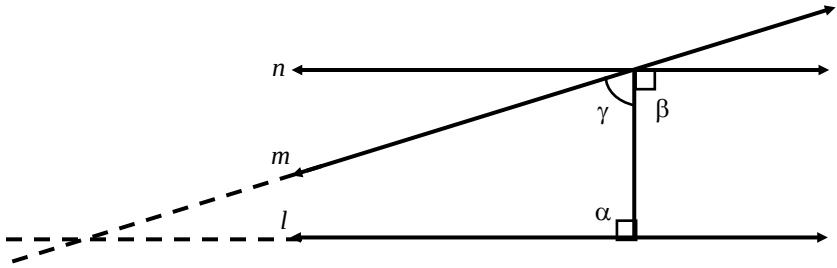
quantity. These govern and constrain, as shown above, the activity of the productive imagination in its construction of basic geometrical objects in finite regions of metaphysical space.

We consider next which system of geometrical objects and propositions can be shown to be *really possible* on the basis of the *real possibility* of the three fundamental descriptions. Since the three fundamental descriptions are comparable with straight-edge-and-compass constructions and since Euclid's system of geometry also captures such processes, it is not surprising that the *real possibility* of Euclid's geometry can indeed be grounded on the three fundamental descriptions. Let us expound on this grounding. Euclid's first three postulates read:

I. *To construct a straight line from any point to any point.*
II. *To construct a finite straight line continuously in a straight line.*
III. *To construct a circle with any center and distance.*

The *real possibility* of geometrical objects constructed according to these three postulates is immediately derivable from the three fundamental descriptions. The first description makes *really possible* Euclid I. Given two points A, B, a third point C constructs a straight line by moving from A to B. The second fundamental description grounds the *real possibility* of Euclid II. Given a straight line *l* and a point A on *l*, a point B extends *l* by a (given) finite straight line moving in the opposite direction of A. Finally, the third fundamental description makes *really possible* Euclid III. Given a center C and a "distance" *r*, a point A describes a circle with ray *r* and center C. Further, Euclid's fifth postulate reads:

V. *That, if a straight line falling on two straight lines makes the interior angles on the same side less than two right angles, the two straight lines, if constructed indefinitely, meet on that side on which are the angles less than two right angles.*



We need to explain the real possibility of a construction in accordance with Euclid V on the basis of the three fundamental descriptions. Thus, consider the diagram above. We need to illustrate that, given that the angle-sum of α and γ is less than two right angles

(i.e. α and the supplement of β), the intersection of lines *l* and *m* can be *constructed* on the side of α and γ in accordance with the three fundamental descriptions. Hence, by description II, *extend* indefinitely lines *l* and *m*. Since the sum of α and γ is not equal to two right angles, lines *l* and *m* do not have *the same direction*, as it were. Since the sum of α and γ is less than two right angles, lines *l* and *m* approximate each other on the side of α and γ. Since the extensions of lines *l* and *m* must, as it were, be *straight* in accordance with description II, it is *really possible* for them to intersect on the side of α and γ; the existence of the point of intersection is in most cases a matter of intuition.[20] This merely illustrates, on the basis of the fundamental descriptions or, equivalently, by a process of straight-edge-and-compass a priori construction, that it is *really possible*, i.e. conceptually and intuitively possible, to construct (or merely conceptually grasp) in pure intuition the point of intersection of any two lines (such as *l* and *m*) under the conditions specified by Euclid V. We shall not discuss the *real possibility* of Euclid's fourth postulate.[21] We thus derive from the three fundamental descriptions the *real possibility* of Euclid's postulates. Geometrical objects constructed in pure intuition in accordance with these postulates and the synthetic a priori propositions derived from these constructions are likewise given *real possibility*.

The above derivation also explains the origin of the *conceptual content* of Kant's pure concepts of Euclid's geometry. The concept of a triangle would contain, for instance, the content of a figure enclosed in three *lines that have unity of direction*, that is, three lines *conceptually* understood to be straight. Such a concept and its content have been first derived from the construction in pure intuition of a triangular figure, whose construction has first shown the *real possibility* of a figure enclosed in three lines. But such construction itself was *conceptually* governed precisely by the *categories of quantity*. Hence, the conceptual content of pure concepts of geometry, that have been constructed and shown to be *really possible*, is derived a priori from the conceptual content of the categories of quantity (i.e. the content of *unity*, *plurality* and *totality*). Only then can the understanding exhibit *in concreto* the pure geometrical concept of, say, the triangle, by again *re*constructing it in pure intuition according to the *schema* specified by the conceptual content in the concept. In addition, the derivation shows the status, in Kant's view, of Euclid's postulates and geometry. They represent a system of geometry – the only possible one for Kant and the only one he was aware of – that is *really possible*, i.e. that agrees with the formal conditions of experience and in which the *intuitive* and *conceptual* features are made possible, respectively, by the a priori forms of sensibility and by the categories of the understanding. This admixture of *intuitive* and *conceptual* features

[20] The conceptual, as opposed to intuitive, grasp of the point of intersection might presumably intervene in cases in which the value of $(\alpha + \gamma)'$ is very close, but not equal to, two right angles.

[21] For some suggestions about Kant's view on "congruence," see *Remark 1* (MFNS, 493).

in Euclid's geometry also suggests a possible explanation for why one who is not acquainted with contemporary, logical treatments of Euclidean geometry is nevertheless wholly convinced of the "a priori necessity" of (most of) Euclid's constructions and proofs in Book I of the *Elements*, as Kant manifestly was.

Section III

We have discussed the extent to which, for Kant, metaphysical space and time as the a priori forms of sensibility ground the possibility of spatiotemporal and kinematic constructions of geometrical objects in pure intuition. Specifically, such faculties have been found to make possible, first, the continuous generation by motion of a point of the geometrical constructions and, secondly, their potentially infinite iterability. Further, we have inquired into the extent to which the categories of the understanding make possible geometrical constructions. We have seen that the categories of quantity govern the activity of the productive imagination in the construction of the three fundamental descriptions of space. On the basis of these three fundamental descriptions, we have derived the *real possibility*, in Kant's sense, of geometrical objects constructed in pure intuition in accordance with Euclid's postulates, thereby deriving the *real possibility* of Euclid's geometry. What we consider next is the question whether Euclid's geometry *de facto* is the only system that can be given *real possibility* in the way expounded in the previous section. That is, given Kant's theory of geometrical constructions as discussed, we ask whether the theory *entails* that no other system other than Euclid's can be *really possible*.

It goes without saying that before J. Bolyai and Lobachevsky made their work on non-Euclidean hyperbolic geometry known to the public in the first half of the nineteenth century, most philosophers and mathematicians conceived of geometry as strictly Euclidean and considered Euclidean geometry to be the correct and only possible description of "the structure of space."[22] Lobachevsky himself initially called his *hyperbolic* system an "imaginary" geometry, presumably in disbelief of the actual existence of any such hyperbolic system.[23] That Kant also was of this opinion stands out clearly from his writings; throughout the *Critique*, geometry is *implicitly* understood to be Euclidean to the extent that geometry is never characterized as such and Euclid is never mentioned or questioned directly. In addition, there is also the fact that non-Euclidean geometries were developed as strictly logical systems, in which theorems are deduced from axioms by logic alone; no room was left for the spatiotemporal and kinematic construction of geometrical objects in pure intuition, the inferential role of intuition, or

[22] The numerous attempts by famous mathematicians, such as Legendre and Farkas Bolyai, to prove the Euclidean parallel postulate from the axioms of neutral geometry are a clear indication of this. See Marvin Greenberg, *Euclidean and Non-Euclidean geometries: development and history*, 4th edition (New York: W. H. Freeman, 2007), ch. V-VI.

[23] Greenberg, *Euclidean and Non-Euclidean geometries: development and history*, ch. VI.

the synthetic nature of geometrical propositions. However, it is important to distinguish what Kant thought of as a natural consequence of mathematical developments in the eighteenth century *from* the actual constraints and implications of his theory. The argument just reproduced, according to which Kant's theory entails that geometry and the geometrical structure of metaphysical space (and time) can only be Euclidean just because non-Euclidean geometries fail to have the synthetic and spatiotemporal character that Kant individuates in Euclid's constructive geometry and adopts in his theory simply misses the point—for it compares a synthetic and constructive treatment of Euclidean geometry with analytic and logical treatments of non-Euclidean geometries. To settle in a reasonable way the issue whether, on Kant's theory, the geometry that is given *real possibility* can be Euclidean, non-Euclidean or both, we must proceed as follows: either (I) by considering *both* Euclidean and non-Euclidean geometries as on their modern, analytic developments or (II) by taking some old-fashioned, constructive-intuitive and synthetic treatments of *both* Euclidean and non-Euclidean geometries. In the former case, we are led to conclude that Kant's theory of geometry is simply outdated. That is, since on their modern developments both Euclidean and non-Euclidean geometries are strictly logical and analytic systems, neither fits the synthetic and spatiotemporal character that Kant individuates in geometry; hence, Kant's theory is not a theory about (modern) geometry at all. However, such argumentation is not helpful to the extent that it assumes a view of geometry that is completely foreign to Kant and that cannot help answering our initial question about Kant's theory, from Kant's perspective. Hence, we are led to consider the latter case (II). We have already discussed the *real possibility* of Euclid's geometry as developed in the *Elements*. What we need to consider next, then, is a treatment of some non-Euclidean geometry analogous to Euclid's treatment of his own geometry. Thus, considering the case of *hyperbolic* geometry, we need to build a treatment of *hyperbolic* geometry analogous to Euclid's and verify whether, on Kant's theory, hyperbolic geometry thusly considered can be a *really possible* system of geometry.

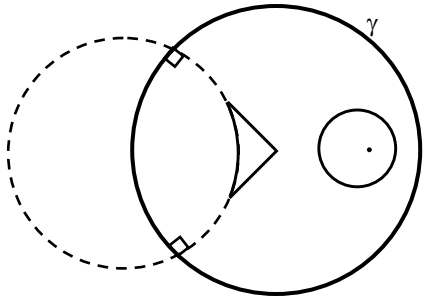
To build a constructive and kinematic treatment of *hyperbolic geometry*, we consider a Poincare disk as our plane and construct in it hyperbolic lines and circles.[24] Specifically, let our *plane* be a Poincare disk, that is, the *interior* of a Euclidean circle γ where the circumference represents merely *ideal points*, and let the following postulates be our constructive-intuitive claims:

I. *To construct a P-line from any point to any point, such that the P-line is either an open diameter of γ or an open arc of a Euclidean circle orthogonal to γ.*

[24] The Poincare disk is, in contemporary developments, an interpretation and model of hyperbolic geometry. We do not consider it here as a model for our constructive-intuitive hyperbolic geometry; rather, as a plane in which to construct hyperbolic objects according to the postulates.

II. *Given any P-segment of a P-line defined as in (I), to continuously extend such P-segment in the P-line.*
III. *For any two distinct points A and B, to construct a P-circle with P-center in A and passing through B.*

To these claims, we could add some P-congruence postulates as well as the hyperbolic parallel postulate. Ideally, these postulates construct P-points, P-lines and P-relations in a Poincare disk in such a way that, if we were to interpret the primitive terms of some modern axiomatization of hyperbolic geometry in terms of these P-points, P-lines and P-relations, we would obtain a model for the hyperbolic axiomatization. Thus, we attain a constructive and kinematic treatment of hyperbolic geometry analogous to a sufficient extent to Euclid's treatment of Euclidean geometry.



The existence of points, lines, figures, and points of intersection between such figures and other similar features implicit in Euclid's treatment are here as well constructed and intuited from the construction, rather than being logically deduced. Hence, we need to verify whether such constructive hyperbolic geometry in the Poincare disk could be made *really possible* on Kant's theory. That is, given the extent to which metaphysical space and time and the categories ground the real possibility of Euclid's system, we ask whether they could ground in a similar way the real possibility of constructive hyperbolic geometry as well. In answering this, it is important to keep in mind that in the Poincare disk all hyperbolic constructions, i.e. the construction of all legal P-lines and P-circles, can be carried out by a precise process of straight-edge-and-compass construction (even though *not all* straight-edge-and-compass constructions in the Poincare disk are constructions of legal hyperbolic P-figures).[25] Thus, our task of verifying the *real possibility* of constructive *hyperbolic* geometry is simplified by the fact that the three fundamental descriptions of metaphysical space identified above suffice, as it

[25] For instance, a Euclidean line in the Poincare disk not passing through the center of the disk is not a legal hyperbolic construction (as our postulate I also claims). For the precise process of construction of hyperbolic figures see Greenberg, *Euclidean and Non-Euclidean geometries: development and history*, Ch. VII.

were, to construct all hyperbolic P-figures in the Poincare disk. That is, since all hyperbolic constructions can be carried out by a process of straight-edge-and-compass construction and since the three fundamental descriptions are equivalent to this process, the three fundamental descriptions suffice to make possible the construction of all hyperbolic P-figures (in pure intuition). Then, on the basis of the real possibility of the three fundamental descriptions of metaphysical space, we derive also the *real possibility* of geometrical objects constructed in pure intuition in accordance with the hyperbolic postulates, i.e. the *real possibility,* on Kant's theory, of constructive hyperbolic geometry. The intuitive features of hyperbolic objects in pure intuition, i.e. their kinematic construction by motion of a point, their continuous generation and their infinite iterability, are grounded in the a priori forms of sensibility, i.e. in metaphysical space and time. Correspondingly, the conceptual features of hyperbolic objects are made possible by the categories, which govern the precise process of straight-edge-and-compass construction through the three fundamental descriptions, thereby governing the construction of hyperbolic P-lines and P-circles. Hence, the boundless intuition of metaphysical space grounds the possibility of constructing in pure intuition any hyperbolic object (of arbitrary finite extension). Metaphysical space and its actual infinity are conceptually *understood* as a hyperbolic plane enclosed in a Euclidean circle of infinite radius and ideal circumference. That is, roughly speaking, since hyperbolic objects are *intuitively* constructed in metaphysical space and time, metaphysical space and time as forms of intuition become *conceptually* understood or conceived of as a *hyperbolic plane* (thereby giving rise to an intuition of space understood to be *hyperbolic*). Accordingly, in such a case, the concept, say, of a hyperbolic triangle shall differ from that of a Euclidean triangle, even though they are both derived from the conceptual features of constructions in accordance with the three fundamental descriptions. The concept of a Euclidean triangle contains, among others, the content of a figure enclosed *in three lines that have unity of direction*. The concept of a hyperbolic triangle, on the other hand, would have the respective content of a figure enclosed in *three lines of which at least one follows a totality of directions.* In this way, then, constructive hyperbolic geometry, i.e. the kinematic and spatiotemporal construction of hyperbolic objects in pure intuition, *can* be grounded upon metaphysical space and time and the categories; that is, constructive hyperbolic geometry can be, on Kant's theory, a *really possible* system along with Euclid's geometry.

From the above analysis, it follows that Kant's theory of geometry is not restricted to Euclidean geometry, but restricted to (any) geometry that comes in a constructive, spatiotemporal and kinematic form determined by constructive-intuitive claims. Metaphysical space and time as the a priori forms of intuition are inherently neither Euclidean nor non-Euclidean; they only ground the possibility of certain *intuitive features* that are conferred upon (any) geometrical objects constructed in pure intuition. Such geometrical objects can be either Euclidean or non-Euclidean in their character; if Euclidean or, respectively, hyperbolic objects are constructed in pure intuition, the concept of metaphysical space as the a priori form of intuition and, hence, its

geometrical structure, is then conceptually conceived of as Euclidean, or hyperbolic. The real possibility of constructing Euclidean or hyperbolic objects in pure intuition not only gives rise to an intuition of space (and time) understood to be Euclidean or hyperbolic (respectively), but also corresponds to the possibility of deriving synthetic a priori propositions about the particular geometry under consideration.

On the conceptual side of Kant's theory, there remains a question about the three fundamental descriptions governed by the categories of the understanding, i.e. whether geometrical objects in pure intuition must be constructible in accordance with the three fundamental descriptions, in order to ensure that the categories of quantity make possible and precisely determine the *conceptual features* of the geometrical constructions. *A fortiori*, it seems reasonable to say that the Kantian categories are not limited to the conceptual grounding of the three fundamental descriptions, i.e. of strictly straight-edge-and-compass constructions. At the same time, however, the understanding seems not to be able to conceptually constrain and grasp the implications of geometrical constructions of very high complexity. Whether the precise limitations of the understanding are represented by the concepts of *unity*, some sort of *plurality* and *totality*, applied in some way to geometrical constructions in pure intuition, is an inquiry that we shall leave open.

Conclusion

In section I, we began by relating Kant's general argument for the *synthetic a priori* character of geometrical propositions; then, focusing on the *construction* of geometrical objects in pure intuition, we related Friedman's approach to Kant, according to which geometrical figures are *spatiotemporal* objects *continuously generated* in the a priori forms of intuition, i.e. space and time. In section II, starting from Friedman's interpretation, we gave an extended account of the construction of geometrical objects in pure intuition. We showed that, on Kant's theory of geometry, both metaphysical space and time as the a priori forms of sensibility *and* the categories of the understanding are involved in grounding the possibility of geometrical constructions in pure intuition. Metaphysical space and time make possible the *intuitive features* of geometrical constructions, that is, their *continuous generation by the motion of a point* and their *potentially infinite iterability*. The categories, on the other hand, make possible the *conceptual features* of the geometrical constructions; that is, by grounding the possibility of the *three fundamental descriptions*, the categories govern and constrain conceptually both the *finite* constructions in pure intuition and, consequently, the constructions that are *understood* to be *potentially infinite*. On the basis of the fundamental descriptions, we have shown the possibility of Euclid's constructive-intuitive geometry and have given an account of how a priori geometrical concepts, such as that of the Euclidean triangle, are *first made possible* and derived a priori from the pure concepts of the understanding, i.e. the categories. Finally, in section III, we turned to non-Euclidean geometries and showed that, notwithstanding a seemingly widespread belief, Kant's theory of geometry *is* compatible

with *non-Euclidean*, or at least *hyperbolic*, constructions in pure intuition. The intuitive features given by metaphysical space and time and the conceptual ones given by the categories are sufficient to ground the possibility of both Euclidean as well as hyperbolic constructions. This approach to Kant's theory of geometry and cognition not only gives an interpretation of Kant's fragments as well as potential insights into the geometer's methodology, but it also explains the possibility of a spatiotemporal intuition not confined to Euclidean space: it seems to be the case indeed that one can quite easily adapt one's spatiotemporal intuition to non-Euclidean geometries.

Works Cited

Allison, Henry. *The Kant–Eberhard controversy*. Baltimore: Johns Hopkins University Press, 1983.

Carson, Emily. "Kant on intuition in geometry," *Canadian Journal of Philosophy* 27:4 (1997), pp. 489-512.

Friedman, Michael. *Kant and the exact sciences*. Cambridge: Harvard University Press, 1992.

----- "Geometry, construction and intuition in Kant and his Successors," in *Between logic and intuition: Essays in honor of Charles Parsons*, edited by G. Sher and R. Tieszen, pp. 186-218. Cambridge: Cambridge University Press, 2000.

Kant, Immanuel. *Metaphysical Foundations of Natural Science*, translated and edited by Michael Friedman. Cambridge: Cambridge University Press, 2004.

----- *The Critique of Pure Reason*, translated and edited by Paul Guyer and Allen Wood. Cambridge: Cambridge University Press, 1998.

Greenberg, Marvin J. *Euclidean and Non-Euclidean geometries: development and history*, 4th edition. New York: W. H. Freeman, 2007.

Posy, Carl J., ed. *Kant's philosophy of mathematics*, New York: Springer-Verlag, 1992.

Shabel, Lisa. *Reflections on Kant's Concept (and Intuition) of Space,* Studies in History and Philosophy of Science 34:1 (2003), pp. 45-47.

Science, Normativity and Knowledge: A (Qualified) Defense of Naturalized Epistemology

BENJAMIN HERSH
Stanford University

Quine's advocacy of naturalized epistemology has left much debate and controversy in its wake. Now that the dust has settled a bit, many consider Jaegwon Kim's polemic response to be the definitive refutation of Quine's view. I think, however, that this issue at heart is significantly murkier than Kim seems to believe, and Kim's treatment of Quine's view of naturalized epistemology leaves much to be desired. In this paper, I lay out Quine's position as explicitly as possible so as to weigh Kim's critique in greater detail. I find that although Kim successfully touches upon the weak points of Quine's mission, his conclusions are too strong.

In the pages to come, I will discuss Quine's epistemological position, as expounded in his provocative paper "Epistemology Naturalized." Characterized very roughly, "naturalized epistemology" constitutes a rejection of traditional, pre-theoretical philosophy and calls for the use of scientific theory to address fundamental epistemic questions. After dissecting Quine's claims in some detail, I will focus on the foremost challenges posed by his outspoken critic Jaegwon Kim in "What is Naturalized Epistemology?" I will contend that although Kim does offer genuine insights about the limitations of naturalized epistemology, his dismissal of it hinges on an unduly narrow-minded approach to the goals of epistemology. In light of this limitation, Kim's arguments against Quine are far less decisive than they may at first appear.

Before introducing Quine's relatively contentious assertions, it would be fruitful to look at his general approach to the business of epistemology. Quine construes traditional epistemology rather broadly, as a search for the foundations of science.[1] Science is envisioned quite broadly as well, so as to include (roughly speaking) the sum of our theoretical knowledge. That is to say, Quine is not merely concerned with particle physics and chemistry, but with mathematics, biology, history, and all other theoretical domains—essentially, every form of knowledge that extends beyond direct empirical observation or sensory impingement. Although Quine shies away from the

[1] W. V. Quine, "Epistemology Naturalized." *Quintessence: Basic Readings from the Philosophy of W.V. Quine*, ed. R. Gibson (Cambridge, MA: Belknap Press, 2004), p. 259.

history of philosophy, his expressed view that epistemology concerns theoretical knowledge as a whole is certainly historically grounded.[2] This is worth noting because, as we will see later, Kim offers a strikingly different view of epistemology's proper scope.

Insofar as epistemology concerns the foundations of science, the notion of a foundation warrants further clarification. Fortunately, Quine offers a rich discussion of this matter. Quine posits a bifurcation within epistemology into a theory of concepts and a theory of doctrine.[3] The conceptual agenda of epistemology consists in putting natural knowledge in terms of sensation and experience—it is essentially an effort to clarify a body of knowledge by putting complex concepts in terms of basic terms. The doctrinal agenda of epistemology, on the other hand, concerns the derivation and justification of natural knowledge in terms of sensory experience. It is assumed that the goal of doctrinal epistemology is to extend the degree of certainty that we attribute to direct sensory experience to the whole of science, and thus make scientific theory as indubitable as the set of observations upon which it is built. These two projects within epistemology are obviously very relevant to one another, but may nonetheless be treated as separate, if mutually beneficial undertakings.

Following the distinction between a theory of concepts and a theory of doctrine, Quine presents a rather succinct history of epistemic thought, depicting a distinct pattern of failure in both conceptual and doctrinal projects. While the details of this historical narrative lie beyond the scope of the present discussion, a few of Quine's points are worth emphasizing. Regarding the conceptual program, every attempt, from Hume to Carnap, ultimately failed.[4] Moreover, even if Carnap's recent attempt had succeeded in allowing us to translate every sentence about the world into observational terms (that is, demonstrating that scientific theory could, in principle, be stated, as a set of propositions about sensory impingements), it would not have entailed the provability of theoretical sentences about the world from only the information gleaned through observation. Therefore, we have little reason to expect Carnap's strategy to actually extend the certainty of direct observations to more abstract theoretical knowledge.

That is, the dismal hopes of completing the doctrinal project would remain virtually unaffected. Quine emphasizes that even simple generalizations about the world

[2] This is vividly illustrated, for instance, in Descartes' metaphorical tree of knowledge—an image which a tacitly shaped much of Europe's intellectual history.

[3] Here, Quine makes a parallel between the reduction of mathematics to logic and set theory.

[4] Carnap, to whom Quine pays special attention, attempted in *Der logische Aufbau der Welt* to systematically reduce natural science to observational terms. This was not a reduction in the sense of restricting the content of science to what is observed, but much closer to the sense in which Russell reduced mathematics to set theory. Although Carnap was certainly not the only thinker to pursue this goal, Quine regards Carnap's efforts as those closest to achieving success.

(e.g. "all swans are white") extend hopelessly beyond direct sensory experience. This presents a logical obstacle in the way of doctrinal epistemology, in that scientific generalizations simply cannot be grounded only in any number of observations.[5] For this reason, it is clear that scientific theory may never be given the same degree of certainty as direct sensory experience, and the doctrinal goals of epistemology are simply beyond hope.

After establishing the pattern of failure in epistemology, Quine directs his efforts to a comparison between traditional epistemology and empirical psychology in terms of their epistemic prospects. Quine's portrayal of psychology as an epistemic alternative may strike some as counterintuitive, and deserves a bit of attention. We can see that, at least insofar as epistemology and psychology both concern the relations between systems of representation (beliefs, theories, etc) and sensory input, they are seemingly quite alike. Although Quine himself did not follow the experimental literature closely,[6] the comparison between epistemology and psychology receives ubiquitous support from the experimental literature. For instance, cognitive psychologists investigate the mechanisms underlying the formation of concepts given a finite set of exemplars or objects belonging to a category, the structure of semantic knowledge, levels of perception, and other aspects of thought that have long been subject to philosophical speculation. However, Quine takes a bold step in asserting that psychology could potentially replace epistemology to a significant degree. This approach has attracted its share of criticism, and will receive further discussion in due course. It is worth noting for the time being, however, that given Quine's broad conception of epistemology, it is consistent for him to treat the two disciplines as covering at least approximately the same ground.

Despite the previously discussed obstacles for traditional epistemology, Quine recognizes that Carnap's ambitious conceptual epistemic project retains some of its appeal. Specifically, if successful, Carnap's project would presumably clarify both the sensory evidence of science as well as the whole of scientific discourse, by reducing all theoretical statements to austere observational, logical, and set-theoretical terms. While neither of these epistemic goals comes close to the Cartesian standard of certainty, they would understandably be of interest. From this point on, Quine's concerns are effectively twofold: to gauge whether these modest conceptual goals of epistemology have any hope of success, and to determine whether empirical psychology would in fact prove to be a more fruitful path of inquiry.

---

[5] This is essentially the problem of induction, although Quine does not put it in these terms here.

[6] I gleaned this from invaluable discussions with Dagfinn Follesdal, Patrick Suppes and George Smith. I do not have direct textual support for the claim, but it certainly explains Quine's vague treatment of "systems of representation."

If successful, Carnap's project would result in a rational reconstruction of scientific discourse in observational terms. It is clear that if one such reconstruction were achieved, it would only be one of many possible constructions. Still, the creation of any such construction would be an arduous feat—and one that has yet to be observed in the history of philosophy. Moreover, such an achievement would amount to little more than an insightful fiction. Empirical psychology, on the other hand, offers the actual account of how cognitive representations and theory form, and would be no less insightful as a descriptive project. Psychology therefore has a distinct advantage, at least in the descriptive domain.

Nonetheless, a rational reconstruction of scientific discourse in observational terms would seemingly offer one valuable benefit that psychology cannot—the translatability of science into the vocabulary of observation, logic, and set theory. Even if a successful reconstruction were treated as a mere fiction, it would grant a special kind of legitimacy to scientific concepts by showing that they could, in principle, be reduced to a conceptually basic level. Psychology simply cannot accomplish this because psychology is concerned with what we actually learn and, as we may easily observe, we do not typically learn concepts in terms of observations and set theory. But even granted that a rational reconstruction promises to hold this advantage over psychology, it is a separate question whether such a construction is actually possible.

Quine emphasizes that, as a matter of fact, even Carnap could not offer genuine translational reduction of scientific discourse to observational terms. In the *Aufbau,* Carnap describes a procedure for mapping sensory information to spatial and temporal coordinates, but this does not constitute an actual method of translation between scientific and sensory terms. In his later writings, Carnap settled for a more liberal standard of translation, using reduction forms. These reduction forms express implications between observational and scientific sentences, but cannot express equivalences. In a sense, they do not offer a true reduction, since scientific concepts could not, in principle, be eliminated by translation. Thus, the project of rational reduction settles for, at most, a non-eliminative degree of translatability. In settling for partial reducibility, it surrenders its promised advantage over psychology.

Having established the failure of Carnap's rational reduction, Quine takes his argument one step further by applying his thesis of indeterminacy of translation. The indeterminacy thesis, as it relates here, states very roughly that for any two languages, there can be no matter of fact in translating sentences of one language into sentences of the other. Of course, there is much to say about indeterminacy, but for my present purposes, I believe that this approximation will suffice. Because rational reduction includes the task of translating sentences from a conceptual, theoretical language into an observational one, it is clear why Quine would perceive indeterminacy as a relevant factor. Indeed, if an epistemologist were to acknowledge that there is no single "correct" translation of scientific concepts into more basic terms, even the most impressive system of reductions would lose its appeal as a solid foundation.

Quine's rejection of reductive epistemology incites relatively little controversy among philosophers. I will now direct my focus to his more provocative remarks. In order to replace traditional epistemology, Quine suggests a new approach entirely. Unlike its predecessor, the new epistemology does not attempt to derive the whole of science, but becomes a part of it. That is, it assumes a scientific perspective and uses the tools of empirical psychology to investigate what had previously been (approximately) the subject matter of traditional epistemology—the connections between sensory input and the construction of theory. Instead of determining what we can know by trying to find observational justification for our beliefs, Quine urges us to investigate the actual psychological mechanisms by which we come to hold our beliefs. Because the traditional project of reducing science to something else is abandoned, there is no serious threat of circularity in taking science for granted. The new, *naturalized* epistemology becomes a scientific investigation of human cognition and systems of representation. In pursuing this line of research, a human subject may come to understand the psychological mechanisms underlying the development of scientific theory itself. Thus, there is a sense in which epistemology is contained within science, but science is contained within a kind of epistemology as well.

Importantly, Quine's "new" epistemology solves otherwise intractable problems about how to understand the nature of psychological phenomena. For example, humans have three-dimensional representations of the world that are derived from two-dimensional patterns of retinal stimulation. The existence of this type of implicit "inference" within perceptual processes would be very troubling if one were still in the business of justifying science on the firmness of observation. This stems from the fact that observations could no longer be treated as discrete or truly foundational, because we would need a standard by which to judge the accuracy of observation. Moreover, profound ambiguities would arise regarding the proper level of representation to treat as relevant to observation. However, the new epistemology easily avoids this problem because it does not assign any special weight or priority to representations of which one is aware. Rather, "awareness" would count as one of many levels of representation, no more important in principle than the others. Quine discusses a variety of cases in which science clarifies or corrects difficulties created by the traditional reliance on naïve notions of perception and cognition. Since the 1960s, the inadequacies of naïve folk psychology have only become more dramatic, as evidenced by phenomena such as change blindness, cognitive priming, semantic dementia, or responsiveness to visual stimuli in cases of cortical blindness.[7] For this reason, one may

[7] In the case of semantic dementia, individuals experience a loss in semantic memory, which occurs both in verbal and non-verbal domains. For instance, a person may lose the ability to recognize or think about living things, but may seem unimpaired in every other way. In cases of cortical blindness, patients appear to be blind and report that they have no visual experience. Yet, they can respond to visual stimuli in an automatic matter, such as catching

regard naturalized epistemology as an especially attractive alternative to traditional approaches.

A few aspects of Quine's reasoning are especially worth mentioning. Particularly, he does not offer a deductive argument for the acceptance of naturalized epistemology. Rather, he demonstrates the faults of one approach and then emphasizes the relative strengths of another, treating the matter not unlike a choice between competing scientific theories—effectively, though an appeal to success. In sum, his argument may be expressed more or less as follows: P1) empirical psychology covers roughly the same ground as traditional epistemology, P2) traditional (reductive) epistemology has never achieved success, P3) empirical psychology is an exceedingly fruitful research program, P4) traditional epistemology cannot offer translational reductions, and therefore offers no special advantage over empirical psychology, Therefore: philosophers ought to abandon the traditional projects of epistemology and turn instead to psychology. This reconstruction, though inexact, captures the main thrust of Quine's line of reasoning.

Given the polemic nature of Quine's claims, it should come as little surprise that his arguments have provoked a flood of critical responses from the philosophical community.[8] Among the throng of replies, however, Jaegwon Kim's often cited "What is Naturalized Epistemology?" arguably stands as the foremost challenge to Quine's epistemic program. I will now outline Kim's critical discussion of naturalized epistemology, as expressed in his influential essay.

From the start, Kim describes epistemology in notably different terms than Quine. In Kim's view, epistemology going back to Descartes has two chief projects: the identification of criteria for the acceptance of beliefs and the application of these criteria to determine what we can be said to know.[9] As a result, the notion at the heart of epistemology is that of the justification of beliefs, and the relevant questions become

---

balls thrown to them. While cognitive neuroscience has a long way to come before we understand such disorders, there is little doubt that it is far better equipped to describe them than the folk-psychological framework that philosophers often adopt. See Dale Purves et al., *Principles of Cognitive Neuroscience* (Sinauer Associates Inc: 2007).

[8] Though not as widely cited as Kim's paper, a number of replies to Quine have emerged, generally along similar lines. See Barry Stroud, "The Significance of Naturalized Epistemology," *Midwest Studies in Philosophy VI*, ed. P.A. French, T.G. Uehling, Jr., and Howard K. Wettstein (Minneapolis: University of Minnesota Press, 1981), pp. 455-471; David Hull, "The Naked Meme," *Learning, Development and Culture,* ed. H.C. Plotkin (London: John Wiley, 1982), pp. 273-327; Fred Dretske, "Perception from an Epistemological Point of View," *Journal of Philosophy* 68 (1971), pp. 584-591.

[9] Jaegwon Kim, "What is Naturalized Epistemology?" *Philosophical Perspectives 2*, ed. James E. Tomberlin (Asascadero, CA: Ridgeview Publishing Co, 1988), pp. 381-406.

ones about the conditions of justification and which beliefs are justified accordingly. The essential requirement of criteria for the justification of beliefs is that their formulation be put in descriptive or naturalistic terms, without the use of evaluative or normative terms of any kind. Thus, terms such as "good reason," "adequate evidence," "having the right to be sure" and so forth must be excluded from the formulation of criteria; they presuppose epistemic normativity and therefore would not provide an informative criterion for epistemic normativity.

Moreover, according to Kim, questions about justification are uniquely important to epistemology. Referring to the traditional depiction of knowledge in terms of justified and true belief, Kim emphasizes that justification is the only strictly epistemic component of knowledge. Belief is a matter of psychology, truth is a matter of semantics and metaphysics, and apparently, it is only through their relation to justification that either gains epistemic relevance. Furthermore, justification makes knowledge an essentially normative concept. By this, Kim means that when one is justified in holding a belief, it is *permissible* to do so, and to hold contradictory beliefs would be *irresponsible*. That is, justification is a matter of which beliefs one *ought* to hold, and which beliefs one ought *not* to hold. To this extent, epistemology is normative in the same way that ethics can be said to be normative.

Next, Kim directs his attention to the target of Quine's critique in "Epistemology Naturalized." Rather than adopting Quine's preferred vocabulary of "traditional epistemology," "rational reconstruction," "eliminative reduction," and so forth, Kim refers broadly to "foundationalist" movements in epistemology. Foundationalist strategies explain justification through a two-step process—the identification of directly justified beliefs, and the justification of other beliefs by means of inference. In this way, the basic, directly justified beliefs serve as an epistemic foundation for all others. Descartes exemplified this approach, regarding beliefs about his immediate mental states (among others) as indubitable, or so obvious that they cannot be doubted. Over time, various philosophers adopted similar assumptions, particularly the tendency to treat beliefs about immediate conscious states as basic. Moreover, some of those who rejected Cartesian mentalism retained his distinctly foundationalist perspective. Kim interprets Quine's attack on "traditional" epistemology as an attempt to target only the Cartesian legacy of foundationalism, and not other non-foundational theories of knowledge.

Regarding Quine's rejection of Cartesian foundationalism, Kim finds little to disagree with. Kim acknowledges the historical failure of both doctrinal and conceptual reductions in foundational epistemology, as well as the implications of Quine's holism[10] for the translatability of scientific concepts into phenomenalistic terms—i.e., that such

[10] Quine's holism underlies much of his work, but an explanation of it lies well beyond the scope of the present paper. What matters in this case is that Kim is willing to acknowledge that Quine's argument seems compatible with his greater theoretical framework.

translations would be impossible. Moreover, Kim also agrees that the quest to "validate" science on the sole basis of sensory experience would be futile, insofar as the process depends entirely on principles of logical deduction. However, Kim takes these points to be rather obvious, perhaps to the point that Quine is only beating a dead horse, so to speak.

When it comes to the interpretation of foundationalism's failure, however, Kim finds serious faults with Quine's reasoning. By Kim's reconstruction, Quine advances two controversial proposals: first, in the wake of Cartesian foundationalism's failure, we should settle for empirical psychology to give us answers about the nature of cognition, and second, empirical psychology may replace the former in such a way that both may be referred to as "epistemology." Immediately, Kim emphasizes that the first proposal is really quite radical. Within the general framework of normative epistemology, a number of approaches may be devised that do not meet the same fate as strict Cartesian foundationalism. For instance, justification may be put in terms of a belief's consistency with a larger belief set, or one might incorporate probabilistic derivation to justify non-basic knowledge. Many such alternatives exist, but in Kim's view, Quine calls for the systematic abandonment of the justification-centered framework and the adoption of a purely descriptive one. Although Quine does explicitly describe epistemology as normative, Kim insists that Quine's implicit agenda is to drive normativity and justification entirely out of the business of epistemology.

Because justification is integral to the notion of knowledge in Kim's view, he asserts that Quine's naturalized epistemology can have nothing to do with knowledge if it abandons justification. It is not clear if Quine would object to this characterization of his position, for as Kim notes, he tends not to speak in terms of "knowledge" at all, but of sensory stimulation giving rise to theories and representations of the world. That is, naturalized epistemology seeks to uncover causal mechanisms underlying cognitive processes and does not, Kim argues, entertain notions of their justification. As far as Kim is concerned, however, the project of naturalized epistemology bears no resemblance at all to that of "traditional" epistemology, which concerns evidential relations above and beyond causal ones. That is, causal relations between "evidence" and "theory" simply are not *evidential* relations in any sense relevant to justification. For instance, knowing that Boyle's observation of mice dying in a vacuum chamber was causally relevant to the development of phlogiston theory hardly indicates whether the said mouse-deaths are sufficient evidence to justify phlogiston theory.

Kim supports this claim about evidential relations by noting that descriptions of causal relations between sensory input and cognitive representations may differ widely across biological species, yet the evidential relations will remain the same (e.g. imagine a silicone-based Martian detective, who rationally forms beliefs on the basis of clues). Kim therefore believes that "traditional" epistemology and naturalized epistemology are involved in fundamentally different types of inquiry, and it is unclear

how one can ever replace the other. Although psychology certainly supplies valuable insights, Kim does not think that it can ever rightfully be referred to as "epistemology."

The core of Kim's line of reasoning against Quine's naturalized epistemology may thus be reconstructed roughly in the following way:[11]

P1) Quine's naturalized epistemology is entirely descriptive
P2) A purely descriptive inquiry is not evaluative
P3) Epistemology is concerned with knowledge
P4) The concept of knowledge is that of true and justified belief
P4) Justification is an evaluative concept
P5) Knowledge has an evaluative dimension
P6) Epistemology must be evaluative
P7) Quine's naturalized epistemology is not evaluative
Therefore: Quine's naturalized epistemology is not epistemology at all

While Kim does not limit his discussion to these points,[12] it is this argument that is commonly taken to be the definitive challenge to Quine's epistemic program.

With regard to the above argument, Kim anticipates a possible objection—that, even if one concedes the distinction between causal and evidential inquiries, naturalized epistemology and epistemology proper nevertheless share the same subject matter of beliefs and representations. If the only significant differences between traditional and naturalized epistemology are that they concern somewhat different properties of beliefs and that one approach is more feasible than the other, then perhaps it may be appropriate to refer to both as "epistemology."

Kim has a response to this objection. Not only is knowledge a normative concept, he asserts, but *belief* is normative as well. His reasoning for this claim takes the following form. When one attributes a belief to a person, one does so with reference to a larger set of beliefs. In order for belief attribution to be meaningful, the cognizer must have a minimal level of rationality, such that his or her cognitive "output" is constrained by rational principles. If the cognizer is apparently irrational, it becomes difficult or impossible to attribute meaningful beliefs to him or her; thus, the principle of charity requires that only rational, or at least consistent beliefs be attributed to a person. For instance, if someone appears to believe both P and not-P simultaneously, it becomes

[11] In this reconstruction, I take a few liberties for the sake of simplicity and clarity. In particular, I do not distinguish between Kim's separate discussions of the "types" of approach and "subject matter," as these would make the reconstruction unduly complicated and obscure the more critical premises.

[12] For instance, he argues that many philosophers who claim to support Quine actually reject his descriptive approach, if implicitly. Additionally, he offers a separate argument for a supervenience account of normativity.

difficult to attribute a meaningful belief at all unless "P" or "not" is reinterpreted to avoid contradiction. Thus, argues Kim, belief attribution requires the evaluation of beliefs. If one wishes to describe beliefs scientifically, one must be able to ascribe them, which presupposes an evaluative procedure. If Quine does in fact reject normative epistemology entirely, he must reject epistemology about beliefs—which would seemingly leave nothing left to call epistemology.

Now that I have outlined the lines of argumentation employed by both Quine and Kim, I may proceed with my own assessment of Kim's criticisms. Before continuing, however, I would like to clarify a semantic issue that motivates some of Kim's criticisms. Let us hearken back to Quine's stated perspective. As a starting point, he regards epistemology as the search for the foundations of science (which is to be interpreted broadly). He discusses two approaches: "traditional" epistemology, which attempts to derive and/or reconstruct science from direct experience, logic and set theory, and "naturalized" epistemology, which is the scientific investigation of how science is formed (empirical psychology). It is well worth noting that Quine does not refer directly to normative justification either in describing the goals of epistemology generally or in describing the "traditional" approach. When he advocates rejecting the project of justification, he refers to the mission to make science as clear and as certain as basic sense data, logic and set theory. Insofar as justification is a distinctly normative question, Quine seems apathetic.[13] Thus, Kim and Quine have very different concepts in mind when they refer to "justification." Nevertheless, this difference in the use of the word should not change the applicability of Kim's arguments.[14] From this point on, I will use the term only in the normative sense.

Now, I will proceed with my critique of Kim's argument. A good starting place would be his explicit reliance on the traditional depiction of knowledge as true, justified belief. While Kim has many precedents for using this definition, it is notoriously imperfect; as Gettier demonstrates in "Is Justified True Belief Knowledge," counter examples can be surprisingly easy to contrive.[15] A typical example of this would be a scenario where a person's watch breaks and displays 12:00 am, and that person happens to look at his or her watch at 12:00 am. In this case, the person would have a true and justified belief that it is 12:00 am, but it seems deeply counterintuitive to attribute knowledge to that person. Although few would respond to such scenarios by denying the need for justification as part of the concept of knowledge, much ink has been spilled, and a generally satisfying solution to Gettier cases that would elucidate necessary

[13] Given the austerity of Quine's ontology, it comes as little surprise that he would avoid committing to special normative properties of knowledge.

[14] When Kim interprets Quine as rejecting normative justification, a charitable reading requires that we look to the "descriptive" nature of naturalized epistemology—not to Quine's explicit rejection of (derivational) justification.

[15] Edmund Gettier, "Is Justified True Belief Knowledge?" *Analysis* 23 (1963), pp. 121-123.

and sufficient conditions for knowledge has yet to be formulated. The lesson that I would like to abstract from this is that even seemingly straightforward matters in epistemology can prove rather elusive. Kim, as it were, scarcely acknowledges the degree of conceptual confusion that plagues epistemology proper, and seems to take it for granted that "proper" epistemologists are in substantial agreement about the nature of their own work.

The power of Kim's main argument comes from his identification of normativity as a criterion for epistemology—specifically, a criterion that naturalized epistemology does not seem to satisfy. At the heart of this claim is Kim's explicit presumption that epistemology itself is *fundamentally about* the evaluation of beliefs. I will demonstrate that Kim's largely unreserved endorsement of this deceptively simple assumption is problematic. Because of this, the root of his argument's polemic strength is also a source of conceptual weakness.

Kim is certainly correct in emphasizing that the evaluation of beliefs plays a significant role in epistemology. One need not read deeply into the historical literature to discover that issues of normative justification figure prominently in the epistemological tradition. Nevertheless, figuring prominently in a tradition is quite different from being a necessary condition for membership in a set. In reality, epistemology is concerned with more than designating which beliefs one ought to accept. Let us examine the tradition of Cartesian foundationalism, to which both Quine and Kim tend to appeal as the paradigm case of traditional epistemology. As far as Kim is concerned, the Cartesian program is motivated by the notion that we are only fully justified in holding beliefs that are certain. Thus, foundational epistemology as Kim describes it may at first seem inseparable from the question of what one *ought* to believe. Nevertheless, this apparent inseparability is contingent on whether one actually *does* regard certainty as a criterion of justification.

I contend that even if one denies a strong connection between certainty and normative justification, epistemic (and even foundationalist) efforts to establish certainty would still be of great philosophical interest. To begin with, I take it to be evident that even if we were to abandon the normative component of foundationalism (especially the notion that we ought to believe all and only propositions that are certain), the project of establishing certainty would hardly seem less philosophically motivated. That is—why wouldn't the topic of certainty be relevant to descriptive theories of knowledge? Why would issues concerting the reliability and veracity of cognitive representations lie beyond the scope of a descriptive project? In fact, reliability and fallibility are central topics in descriptive work in empirical psychology, and are central notions in describing how the brain interacts with an organism's environment. Given these considerations, it seems clear the tie between certainty and normativity is not an absolute one, and at most a matter of degree.

Moreover, one can ignore the issue of certainty altogether and still find substantial philosophical value in a foundationalist project; as Quine emphasizes, many

foundationalist (reductionist) projects are highly motivated by the desire to clarify theoretical knowledge. For the sake of convenience, I will refer to the class of epistemic projects that do not concern normative justification as "descriptive."

Consider two hypothetical foundationalists, Lyle and Téodor. Both philosophers spend years trying to deduce their theoretical knowledge on the basis of immediate experience, logic and set theory. Lyle does this because he wants to know which of his beliefs he is justified in believing, whereas Téodor does it with the hope of gaining descriptive insights into the nature of his theoretical knowledge. Suppose that Lyle and Téodor use the exact same methods of inquiry and come to identical conclusions about the extent of provability. It would seem strange to say that Lyle is doing epistemology and that Téodor is not. Rather, it would be more appropriate to say that they were both pursuing epistemological projects but with different goals in mind. I think that this should be sufficient grounds to doubt Kim's insistence that epistemology is *necessarily* concerned with justification.

I will concede that Kim is correct in denying the ability of naturalized epistemology to replace foundationalism to the extent that epistemology is a normative project. However, naturalized epistemology seems perfectly capable of replacing the descriptive projects of foundationalism, which are not properly acknowledged by Kim. While naturalized epistemology cannot replace all of the projects included under the umbrella of foundationalism, there is no reason that one cannot replace the subset of distinctly descriptive projects about the clarification and derivation of theory with empirical psychology. Moreover, much work in recent "traditional" epistemology (e.g. Dretske's fully penetrating operators[16] or Lewis's contextualism[17]) scarcely does justice to epistemology's historical goal of clarifying the structure of perception and knowledge. If we were to hold these approaches to a similar standard of continuity with previous epistemic projects as Kim does with regard to normative justification, we may be tempted to decide that they are not genuine epistemology either. Presumably, Kim does not wish to do this. Thus, Quine's naturalized epistemology can replace "traditional" foundationalist epistemology, in a limited but very significant sense.

At this point, I would like to address one possible response to my argument thus far: biting the bullet. I understand the above discussion is unlikely to dissuade someone who already firmly believes that epistemology must concern notions of justification. With regard to the Lyle and Téodor example, such a person can always bite the bullet and maintain that Lyle engages in epistemic inquiry but not Téodor. I think that the proper response is simply to emphasize that a term does not require necessary or sufficient conditions for its use. I wonder if perhaps the desire to preserve normativity as a necessary condition is tied to an implicit essentialist intuition to the

[16] Fred Dretske, "Epistemic Operators," *Journal of Philosophy* 67 (1970), pp. 1007-1023.
[17] David Lewis, "Elusive Knowledge," *Australasian Journal of Philosophy* 74:4 (1996), pp. 549-567.

effect that any category *must* be defined through a set of core features. While a handful of contemporary philosophers may find essentialism palatable in certain contexts,[18] its shortcomings may be illustrated with the simple (yet seemingly impossible) task of providing a clear set of sufficient and necessary conditions for an activity to qualify as a game or an object to qualify as a chair.[19] It might be more appropriate to think of the term "epistemology" as applying to various historical traditions of inquiry that are bound by sets of overlapping features, none of which is common to all— that is, Wittgenstein's notion of "family resemblances."[20] This approach could acknowledge the importance of justification in epistemology without treating it as an essential trait. Moreover, naturalized epistemology could fit neatly under the umbrella of epistemology on account of its similarities to certain members of the category, yet remain relatively dissimilar to others. This is by no means the only approach that one might take in delimiting epistemology, but it strikes me as a good deal more capable of accurately capturing the category than essentialism.

Now, I would like to address Kim's second challenge to naturalized epistemology: that it must presuppose normativity in order to implement a science of beliefs. As I discussed previously, Kim asserts that belief attribution requires that beliefs be evaluated for a minimal degree of rationality. He thinks that because Quine's naturalized epistemology is intended to be purely descriptive, this will be a major problem; it must either presuppose evaluative procedures for belief attribution or cease to be concerned with beliefs at all.

The polemic force of Kim's objection stems from an instance of equivocation in his interpretation of naturalized epistemology, which leads him to portray its "descriptive" agenda in a misleading fashion. Kim initially characterizes naturalized epistemology as descriptive in the sense that it does not seek to evaluate the cognitive representations that it describes. In this sense, evaluation is a matter of normative justification. However, Kim uses a significantly different notion of evaluation when describing the process of belief attribution. I would contend that belief attribution does not require that one ask whether a cognizer's beliefs are justified; it only requires that one translate his or her beliefs in a systematic way that avoids the attribution of

[18] A noteworthy example would be Saul Kripke, whose reasoning stems from modal logic. It may be worthwhile to note that Quine himself saw essentialism as a reductio of extensional applications of modal logic, warning that quantification in modal contexts "leads us back into the metaphysical jungle of Aristotelian essentialism" in the final sentence of "Three Grades of Modal Involvement" (1953), in *Quintessence: Basic Readings from the Philosophy of W.V. Quine*, edited by Gibson, R. Cambridge, MA: Belknap Press, 2004, p. 259.

[19] Insofar as "epistemology" is to be understood as term in natural language or a psychological concept, essentialism is vastly and obviously inadequate.

[20] Ludwig Wittgenstein, *Philosophical Investigations.* Trans. Anscombe, G.E.M. (Malden, MA: Blackwell Publishing, 1953). Part I, section 67.

inconsistent beliefs. The normativity that is being "presupposed" is not at the level of scientific description, but at the level of scientific procedure. Procedural normativity, however, is nothing unusual in science; every theoretical field has some set of procedures for the evaluation of evidence, be it counting tree rings in a certain order or running one or another method of statistical analysis. The fact that a psychologist "evaluates" the rationality of beliefs does not make his or her scientific descriptions any more normative than those of the hard sciences.

Kim may nevertheless insist that in fact, the normativity involved in belief attribution also exists at the level of description because the scientist would say that the beliefs themselves are rational. However, I would respond that rationality in this sense does not need to be interpreted as normative at all. In attributing a rational belief to a cognizer, the scientist acknowledges that the belief meets a minimal standard of consistency with the other beliefs, but this fact alone does not mean that the cognizer *should* hold that belief. After all, a minimally rational belief set may turn out to be false, and the scientist may well have more rigorous standards of justification than minimal rationality. For this reason, I think that it is misleading for Kim to describe the evaluative aspects of belief attribution as a challenge to naturalized epistemology.

I will now bring my critical discussion to a close. While it should be clear by now that I am not won over by Kim's arguments, I think that he does offer important insights about the limitations faced by Quine's epistemic approach. Most notably, Quine ignores the issue of justification, and in so doing his discussion overlooks much of the Western epistemological tradition. Nevertheless, Kim moves too far in the opposite direction by characterizing epistemology as a single-minded mission for criteria of justification. Ultimately, I believe that Kim's argument against the epistemological relevance of naturalized epistemology is undermined by this tendency. The question on the lips of the philosophical community should not be whether naturalized epistemology could replace its predecessors *qua epistemology* (which I maintain it can, to the vast extent that it shares a family resemblance with its historical relatives), but whether it offers a deeper level of understanding than previous approaches. I think that the answer is clear. Although empirical psychology and cognitive neuroscience are still in their infancy, the last century has seen tremendous strides towards an unprecedented, mathematically precise, rigorously corroborated and comprehensive theory of knowledge. Scientific theory is a powerful tool that allows us to overturn or clarify the vague pre-theoretical notions that plague folk psychology and lie at the heart of a largely stagnant philosophical tradition. Epistemologists have no excuse not to use it, and Quine was certainly "justified" to urge on the philosophical community.

Works Cited

Carnap, Rudolf. *Der Logische Aufbau der Welt*. Leipzig: Felix Meiner Verlag, 1928. English translation by George, Rolf A. *The Logical Structure of the World and Pseudoproblems in Philosophy*. California: University of California Press, 1967.

Descartes, Réné. *Principles of Philosophy*. Translated by V.R. Miller and R.P. Miller. Dordrecht: D. Reidel, 1983.

Dretske, Fred. "Perception From An Epistemological Point of View." *Journal of Philosophy* 68 (1971), pp. 584-591.

----- "Epistemic Operators," *Journal of Philosophy* 67 (1970), pp. 1007-1023.

Gettier, Edmund. "Is Justified True Belief Knowledge?" *Analysis* 23 (1963), pp. 121-123.

Hull, David. "The Naked Meme." *Learning, Development and Culture*, edited by H.C. Plotkin. London: John Wiley, 1982, pp. 273-327.

Kripke, Saul. "Naming and Necessity." *Semantics of Natural Language*, edited by D. Davidson and G. Harman. Dordrecht and Boston: Reidel, 1972.

Lewis, David. "Elusive Knowledge."*Australasian Journal of Philosophy*, 74:4 (1996), pp. 549-567.

Kim, Jaegwon. "What is Naturalized Epistemology?" *Philosophical Perspectives 2*, edited by James E. Tomberlin. Asascadero, CA: Ridgeview Publishing Co, 1988, pp. 381-406.

Purves, Dale et al. *Principles of Cognitive Neuroscience*. Sinauer Associates Inc.: 2007.

Quine, W.V. "Epistemology Naturalized" and "Three Grades of Modal Involvement." *Quintessence: Basic Readings from the Philosophy of W.V. Quine*, edited by Gibson, R. Cambridge, MA: Belknap Press, 2004.

Stroud, Barry. "The Significance of Naturalized Epistemology." *Midwest Studies in Philosophy VI*, edited by P.A. French, T.G. Uehling, Jr., and Howard K. Wettstein. Minneapolis: University of Minnesota Press, 1981, pp. 455-471.

Wittgenstein, Ludwig. *Philosophical Investigations*. Trans. Anscombe, G.E.M., Malden, MA: Blackwell Publishing,1953.

Consequentialism and Rights

PUNEET DHALIWAL
University of Warwick

Consequentialism is often criticized on the grounds that it justifies the gross violation of individual rights in order to bring about the best overall consequences. In this paper I argue that such criticism is mistaken and that consequentialism is able to accommodate respect for rights. Consequentialism does not necessarily demand the violation of rights, if rights are understood in a positive sense as capabilities and the ability to realize important goals, instead of merely the negative demand for non-interference from other moral agents. Moreover, consequentialists do not have to view rights as mere conduits to social utility; they can acknowledge the importance of preserving individual rights while taking into account the social context in which rights are to be protected.

It is a common objection to consequentialist ethical theories, particularly to utilitarianism, that they make demands that are intuitively unpalatable, requiring the violation of individuals' rights. Notwithstanding the obvious difficulties of evaluating so broad a school of ethical theories as consequentialism and so contested a concept as rights, I argue that consequentialism does *not* fail to respect individual rights. To show this, I begin by considering the meaning of consequentialism, in a bid to extract the essence of those moral theories that make normative judgments on the basis of consequences. I then argue for a "unified" approach to individual rights that eschews the extremes of both deontological, constraint-based rights and purely instrumentalist views of rights. Such an approach, I claim, enables consequentialism to respect rights.

Consequentialist Theories

"Consequentialism" refers to a broad school of ethical theories rather than one single doctrine. Consequentialist theories may be direct (they may evaluate the consequences of particular acts) or indirect (they may evaluate the consequences of various rules, motives or virtues that govern acts). They may require the maximising of some good, like happiness in the case of classical

utilitarianism, or they may be content with a "satisficing"[1] approach, which demands only that a certain level of good (however defined) is attained. Further, consequentialist theories may dispute what actually constitutes a good consequence.

Consequentialism is perhaps best understood in opposition to deontology, which normatively categorizes acts as intrinsically good or bad based on particular moral duties. For instance, Kant, through his formulations of the categorical imperative, argued that we, as rational, moral agents, always have a duty to refrain from committing suicide or from making false promises.[2] For the deontologist, therefore, certain actions can be categorized as morally right or wrong, *whatever their consequences.*[3] In contrast, for consequentialism, the virtue of action depends *only* on the consequences. No actions are categorically forbidden or required by appeal to certain metaphysical claims such as inviolable natural rights.[4] The moral content of actions is simply contingent upon their real consequences in the world. Hence, "the right act in any situation is the one that will produce the best overall outcome, as judged from an impersonal point of view."[5]

The "Rights Objection"

If the normative justification for actions comes only from their consequences, then a violation of rights[6] that brings about "better consequences" is presumably

[1] "Satisficing" is a word formed by blending "satisfying" and "sufficing." For an outline of satisficing consequentialism and a defence of such theories, see Michael Slote, "Satisficing Consequentialism," *Proceedings of the Aristotelian Society, Supplementary Volumes*, 58 (1984), pp. 139-63.

[2] Immanuel Kant, *Groundwork of the Metaphysics of Morals*, trans. and ed. M. Gregor. Cambridge: Cambridge University Press, 1998, pp. 31-32 (AK 4:422).

[3] For such a view see G.E.M. Anscombe, "Modern Moral Philosophy." *Philosophy*, 33:124 (1958), pp. 1-19. Anscombe, in arguing for deontology and against consequentialism, warned against being "tempted by fear or hope of consequences" (p. 11).

[4] While deontological theories do not necessarily presuppose certain metaethical positions, certain stock deontological concepts, such as natural rights, clearly sit more comfortably within certain metaethical frameworks, particularly moral realist ones, in which at least some ethical propositions are true in an objective sense.

[5] Samuel Scheffler, *The Rejection of Consequentialism: A Philosophical Investigation of the Considerations Underlying Rival Moral Conceptions* (Oxford: Clarendon Press, 1994), p. 167. Of course, more complex consequentialist theories may allow room for agent-relative concerns such that the notion of the "impersonal point of view" is relaxed somewhat. However, it is the more "impersonal" consequentialist theories that concern us here, for they present the greatest potential for a conflict with personal values and individual rights.

[6] There might be a moral difference between a violation of rights that comes about as a side effect of the action that leads to good consequences, or as a clear intended means by which the good consequences are generated. Hence the Doctrine of Double Effect, from the

justified. Consider, for example, McCloskey's Sheriff Case. A sheriff of a small town must make a choice between two evils. An awful crime has been committed, and the public will violently riot if justice is not meted out soon. There are no real suspects, however, and the only way to prevent rioting is to frame and execute an innocent person. The sheriff must choose between doing nothing and the scapegoat.[7] Assuming that the violence of the ensuing riot is a worse outcome than the execution of one innocent person, the rights objection argues that consequentialism logically demands the execution of the innocent person and cannot, therefore, be considered to respect individual rights.

Viewed properly, however, consequentialism does respect rights. More specifically, the "unified approach" to rights shows how one can simultaneously pursue the best consequences and respect rights.

Rights: A Unified Approach

Many versions of the "rights objection" to consequentialism presuppose a *negative*, deontological, constraint-based view of rights, along the lines of Nozick's "side constraints." For Nozick, rights are constraints that set strict limits on what one may rightfully do to others.[8] It is not at all clear, however, that rights have to be absolute constraints in order to be taken seriously. From Hume we can learn that rights are best conceived of as broad social conventions that are better justified with an eye towards their social utility rather than with any higher metaphysical claims. In his discussion of the origin of justice and property, Hume writes: "A man's property is some object related to him. This relation is not natural, but moral, and founded on justice."[9] He therefore views rights, such as property rights, as "human conventions" based on constructed "artifices." In like manner, I hold that natural rights, as some inviolable metaphysical reality, are indeed, as Bentham remarked, "nonsense upon stilts."[10] Moral rights are better understood as *socially* constituted and tangibly beneficial *social* rules.

Thomas Scanlon, although himself critical of some aspects of many consequentialist theories – for instance, the utilitarian aim of maximising aggregate

---

discussion on the Trolley Problem; see Phillippa Foot, *Virtues and Vices and Other Essays in Moral Philosophy* (Oxford: Blackwell, 1978). It is clearly debatable as to whether consequentialism might demand the "violation of rights" as a means to generate good consequences or merely as a side effect of such actions.

[7] H. J. McCloskey, "An Examination of Restricted Utilitarianism." *The Philosophical Review*, 66:4 (1957), pp. 468-469.

[8] Robert Nozick, *Anarchy, State, and Utopia* (Oxford: Blackwell, 1974), p. 29.

[9] David Hume. *A Treatise of Human Nature*, ed. D. F. Norton and M. J. Norton (Oxford: Oxford University Press, 2000), p. 315.

[10] Jeremy Bentham, "Anarchical Fallacies," in *The Works of Jeremy Bentham, Vol. 2*, ed. J. Bowring (New York: Russell & Russell Inc., 1962), p. 501.

happiness – notes that "rights themselves need to be justified somehow, and how other than by appeal to the human interests their recognition promotes and protects? This seems to be the incontrovertible insight of the classical utilitarians."[11] The utilitarian J.J.C. Smart argues that "the purpose of morality is to subserve the general happiness,"[12] so that moral rights can be conceptualised as means by which the general happiness is promoted. I would not subscribe to a full utilitarian view like Smart's. Though I would agree that moral concepts, such as rights, ought to be strongly related to social welfare, not mere private egoism,[13] I regard purely instrumentalist views of rights as inadequate, as they greatly diminish the normative force of rights by treating them as little more than mere conduits through which utility is generated.

Note that in this paper, I do not attempt to *prove* that rights have no absolute metaphysical foundation. Rather, I seek to show that there is a plausible conception of rights – one that understands rights as a set of social rules instead of inviolable metaphysical absolutes – that enables consequentialism to take rights seriously.

I call this conception of rights a "unified" approach because it captures both the instrumental social value and the normative force of rights in protecting individuals. Individual rights, instead of consisting of merely *negative* constraints on the actions of others, also encompass a *positive* aspect. That is, rights should be conceptualised in terms of individuals' positive claims to entitlements, personal capabilities and adequate opportunities for self-realization. Rights therefore incorporate both a positive concern for self-realization and social welfare, in addition to the absence of interference or harm.

An example that helps to construct a "unified" conception of rights is Amartya Sen's "goal rights systems" account of rights, including his "capabilities" perspective. "Goal rights systems" take into consideration the "fulfilment and non-realization of rights" in the evaluation of states of affairs. That is, rights are considered, at least to some extent, as goals to be promoted rather than Nozickean constraints to be respected absolutely. Sen makes the case for his "goal rights system" with his case of Ali and the Bashers. He illustrates the problem of a potential attack by a gang of "Bashers" on Ali,

---

[11] Thomas M. Scanlon, "Rights, Goals, and Fairness." In *Consequentialism and its Critics*, ed. S. Scheffler (Oxford: Oxford University Press, 1988), p. 74.

[12] J.J.C. Smart, "An outline of a system of utilitarian ethics," in *Utilitarianism, For and Against*, J.J.C. Smart and Bernard Williams (London: Cambridge University Press, 1973), p. 68.

[13] It might be objected, by an argument paralleling Adam Smith's "invisible hand" that rights that strictly (and only) protect private interest would result in a higher aggregate social welfare. See Adam Smith, *An Inquiry into the Nature and Causes of the Wealth of Nations*, ed. K. Sutherland (Oxford: Oxford University Press, 1993), p. 292. While individual rights that engender a healthy private sphere may well contribute to social welfare, they also need to be complemented by a more direct and collective regard to social welfare. Otherwise, protection of private interest alone could yield substantive social inequities that directly harm social welfare.

who has gone out and has no awareness of the danger he faces. His friend, Donna, does not know Ali's whereabouts but knows that there is a note with that information in the room of their friend, Charles. Donna must decide whether or not to violate Charles's privacy by breaking into his room in order to retrieve the note, thereby allowing her to save Ali.

Sen uses this example to highlight the inadequacies of both purely welfarist consequentialist approaches and constraint-based deontology. Welfarist consequentialism, which evaluates consequences based on their effect on welfare, could conceivably justify the bashing of Ali if the bashers derive enough welfare from the attack to outweigh Ali's loss of welfare. If rights are merely tools to promote welfare, Ali's right to be free from the attack can be outweighed by the bashers' greater welfare. Hence a purely welfarist instrumentalist view of rights will not do if we want to save Ali. The Bashers example also shows the limited usefulness of deontological side-constraints, because obeying such side-constraints would prevent Donna from violating Charles's right to privacy by looking for the note in Charles's room. Again, this will not do if we want to save Ali.

In order to save Ali, Sen argues, we must adopt a consequentialist and non-welfarist approach, which can be achieved with his goal rights system. The violation of Charles's right to privacy can be justified on the grounds that it averts a far greater rights violation, namely, the bashing of Ali. In other words, one can perform a consequentialist calculation that takes the rights of both Ali and Charles into consideration when evaluating the different states of affairs. Since the consequences of violating Ali's right are so much worse than the consequences of violating Charles' right, one comes to the conclusion to violate Charles' right and to save Ali.

Sen's use of the "capabilities" perspective to incorporate goal rights into the evaluation of states of affairs further helps us to achieve an understanding of rights that is compatible with consequentialism. Sen conceptualizes goal rights as positive "capabilities" to which one has a right, rather than some relation between two parties involving some particular (and generally negative) correlative duty of non-interference. In an extension of Sen's thought experiment, let us say that Ali faces damage from a natural disaster of which Donna has knowledge but Ali does not. Donna must decide whether or not to break into Charles's room to take his car-keys in order to drive his car to save Ali. In this scenario, the only violation of rights in the traditional, negative, deontological sense is of Charles's right to privacy. Ali's rights are not in danger of being violated in the negative deontological sense, since the damages suffered by Ali would have no agency associated with them. However, if we conceptualize rights in a more positive way by means of capabilities, we see that Ali may potentially suffer damage to his rights in this scenario too. Ali's rights, if they are taken to include "the capability of

moving about without harm,"[14] will be violated if Donna does not infringe Charles's right to privacy in order to save Ali. Since the negative social consequences of violating Ali's right are much greater than the consequences of violating Charles's right, Donna can be justified in infringing Charles's right to privacy in order to save Ali. If we conceptualize rights as capabilities instead of as non-interference requirements binding solely on moral agents, then, we are better able to explain the necessity for Donna to save Ali.

Sen's goal-rights system and the capabilities approach combine what I take to be the individualistic and social faces of rights. The rights themselves, of course, are still individualistic insofar as they pertain to *individuals*: individuals have the right to the capacity to live securely and free from harm. But as rights under the capabilities approach are more positive and assertive than negative deontological rights, respect for rights is far more demanding than mere non-interference with individuals. Respect for rights under the capabilities approach entails the protection and furtherance of the conditions necessary for individuals to enjoy particular capabilities. It may well require a more vigorous form of moral engagement by individuals like Donna, who are not directly responsible for Ali's potential bashing but who are conceived as relevant members of a moral community – that is, a community of moral agents. If we want to live in communities with people who are obligated to do more for one another than refrain from interfering with one another's lives, then we are better off with a "unified" conception of rights that takes into consideration both their significance for individuals and the social context in which rights are protected.

Consequentialism and Rights

With this conception of rights in mind, I argue that consequentialism does respect individual rights. First, consequentialism does not necessarily violate individual rights even in the traditional deontological sense. Second, and most importantly, I will argue that even in those extreme cases in which consequentialism might demand action that violates deontological rights, the demands of consequentialism are compatible with the "unified" conception of rights.

Consequentialism would not routinely demand outcomes such as the execution of the innocent. In most cases, abiding by traditional deontological rights would likely produce the best consequences. In the Sheriff Case, for instance, the consequentialist could oppose the execution of the innocent scapegoat on the grounds that it would likely have substantial detrimental effects on trust in the justice system and societal security. Nevertheless, one might "trap" the consequentialist in some "repugnant conclusion" that demands the violation of individual rights. One might add to the Sheriff Case, as McCloskey does, the provisos that no one need ever discover the

[14] Amartya Sen, "Rights and Agency," in *Consequentialism and its Critics*, ed. S. Scheffler (Oxford: Oxford University Press, 1988), p. 201.

framing of an innocent person.[15] It seems that even though consequentialism may well respect individual rights in "normal" circumstances, in extreme instances, better consequences clearly *would* follow from a situation in which rights are violated in the traditional deontological sense.

Consequentialism, then, would not entirely rule out outcomes that violate deontological rights. Indeed, in extraordinary circumstances, consequentialism may demand the sacrifice of one innocent to save thousands from certain and terrible death; that is, rights would be "violated" to prevent moral catastrophe. However, based on the "unified" conception of rights, I argue that such an outcome is not necessarily a violation of rights.

Rights under the "unified" conception are not absolute metaphysical realities demanding non-interference regardless of the circumstances, but rather a set of valuable socially constructed rules. As social conditions change, the appropriate rules and rights should be sensitive and responsive to such alterations. Moreover, rights are best understood not as negative deontological constraints but as positive capabilities and the ability to realize one's goals. The imperative of preserving the basic stability of a community could, in certain circumstances, take priority over the security of one person, precisely because of the rights involved. The members of the community each have rights to exercise their capacities and realize their goals, and the entire community's ability to function is a prerequisite for the fulfilment of these rights. If the riots in the Sheriff Case, would disrupt the basic functioning of the community, then failing to frame a person who did not commit the crime would lead to a significant rights violation.

Of course, the innocent person has rights under the "unified" conception as well. According to some versions of consequentialism, which evaluate consequences based at least partially on considerations of justice, the rights of the innocent person might even carry much more weight than the rights of other members of the community. The point is merely that the rights of the innocent person should not be seen as utterly inviolable. At some point, they may justifiably be outweighed by the prospect of widespread suffering and harm, especially the kind of suffering that would lead to a breakdown in the legal system and undermine the very possibility of justice in the community. Under the "unified" conception of rights, then, particular rights would sometimes be violated in order to prevent wider and more general violations of rights. The consequentialist, even when making such demands, is still holding a position that is compatible with respect for rights. Recognizing the necessity of trade-offs among the rights of individual members of a society – just as there are trade-offs in the case of Ali and the Bashers – does not imply a lack of concern for rights.

[15] Of course, under such provisos, Benthamite utilitarianism may well demand the use of torture in private, but that is not necessarily representative of all consequentialism.

The deontologist who refuses to sacrifice one individual to save thousands on the grounds of the inviolability of individual rights is, I think, culpable for the deaths of thousands; he or she cannot claim some "moral distance" from the suffering of others in the moral community simply because he or she has complied with minimal conditions of non-interference. Placing rights on some pedestal as an abstract, inviolable concept with little social applicability diminishes rather than respects them. It is only by acting with regard to the social utility of rights, but without reducing them to mere conduits of utility, that they can be taken seriously as a moral concept.

One might object, however, that my argument is predicated on a flawed conception of individual rights in that it is overly collectivist and does not take seriously the *separateness of persons*. John Rawls argued against utilitarianism that it "does not take seriously the distinction between persons,"[16] thus highlighting concerns with consequentialism's tendency to evaluate decisions based on aggregate outcomes. I would reply that consequentialism, under my understanding of rights, takes very seriously individuals' *distinctness* and *autonomy*, and sees rights associated with autonomy as goods to be promoted when evaluating states of affairs. But while consequentialism recognizes individuals' autonomy, it also notes their social constitution and embeddedness, and the ethical importance of the social context of rights. Those who exalt the separateness of persons fail to take seriously the *impersonality of separateness*, which would yield an egoistic and tenuous moral community.

To conclude, I have argued that consequentialism does not require the violation of rights, properly understood. Under a conception of rights that incorporates the positive goals of self-realization and the use of human capabilities, instead of relying solely on the requirement of non-interference, consequentialism can accommodate respect for rights. Consequentialism can provide us with a way to realize rights in the social world, in which we must take into consideration both the significance of individual goals and the social consequences of individual action.

Works Cited

Anscombe, G.E.M. 1958. Modern Moral Philosophy. *Philosophy*, 33 (124), pp. 1-19.

Bentham, Jeremy. "Anarchical Fallacies." *The Works of Jeremy Bentham, Vol. 2*, edited by J. Bowring [1843]. New York: Russell & Russell Inc., 1962.

[16] John Rawls, *A Theory of Justice*. (Cambridge, MA: Harvard University Press, 1971), p. 27.

Foot, Phillippa. *Virtues and Vices and Other Essays in Moral Philosophy*. Oxford: Blackwell, 1978.

Hume, David. *A Treatise of Human Nature*, edited by D. F. Norton & M.J. Norton. Oxford: Oxford University Press, 2000.

Kant, Immanuel. *Groundwork of the Metaphysics of Morals*. Translated and edited by M. Gregor. Cambridge: Cambridge University Press, 1998.

McCloskey, H. J. "An Examination of Restricted Utilitarianism." *The Philosophical Review*, 66:4 (1957), pp. 466-485.

Nozick, Robert. *Anarchy, State and Utopia.* Oxford: Blackwell, 1974.

Orend, Brian. *Human Rights: Concept and Context*. Canada: Broadview Press, 2002.

Pettit, Pettit. "Consequentialism." *Consequentialism*, edited by Stephen Darwall. Dartmouth: Aldershot, 1993, pp 95-107.

Rawls, John. *A Theory of Justice.* Cambridge, MA: Harvard University Press, 1971.

Sen, Amartya. "Rights and Agency" [1982].*Consequentialism and its Critics*, edited by S. Scheffler. Oxford: Oxford University Press, 1988, Ch. 9.

Scanlon, Thomas M. "Rights, Goals, and Fairness" [1978]. *Consequentialism and its Critics*, edited by S. Scheffler. Oxford: Oxford University Press, 1988. Ch. 4.

Scheffler, Samuel. *The Rejection of Consequentialism: A Philosophical Investigation of the Considerations Underlying Rival Moral Conceptions*. Oxford: Clarendon Press, 1994.

Sinnot-Armstrong, Walter. Consequentialism. *The Stanford Encyclopedia of Philosophy* online, edited by E. N. Zalta (updated Spring 2007).

Slote, Michael. Satisficing Consequentialism. *Proceedings of the Aristotelian Society, Supplementary Volumes*, 58 (1984), pp. 139-63.

Smart, J.J.C. & Williams, Bernard. *Utilitarianism, For and Against*. London: Cambridge University Press, 1973.

Smith, Adam. *An Inquiry into the Nature and Causes of the Wealth of Nations*. Edited by K. Sutherland. Oxford: Oxford University Press, 1993.

Sumner, L. W. *The Moral Foundation of Rights*. Oxford: Clarendon Press, 1987.

Thomson, Judith J. The Trolley Problem. *The Yale Law Journal*, 94:6 (1985), pp. 1395-1415.

Interview with Hilary Putnam, Harvard University

RACHEL BAYEFSKY and ERIN MILLER
Yale University

Hilary Putnam is Cogan University Professor Emeritus at Harvard University. He has developed a reputation for excellence in many areas of philosophy, including epistemology, metaphysics, and the philosophy of mind, science and language. A sampling of his books include *Philosophy of Mathematics: Selected Readings* (1964), *Philosophical Papers* (1975 and 1983), *Reason, Truth, and History* (1981), *Pragmatism: An Open Question* (1995), *The Collapse of the Fact/Value Dichotomy and Other Essays* (2002), *Ethics without Ontology* (2004), and *Jewish Philosophy as a Guide to Life* (2008). This interview was conducted at Harvard on June 15, 2009.

Yale Philosophy Review: *You once held a largely scientific materialist worldview. But now you reject science as an appropriate model for philosophical inquiry and encourage philosophers to consider the irreducible ethical and social dimensions of human life. What prompted this shift?*

Hilary Putnam: I'm really puzzled by this description of my intellectual history. I am a scientist in a sense—I am a mathematician. In 1960 I had tenure in what was at the time perhaps the best mathematics department in the world, at Princeton University, and I had and still have a great deal of respect for science. But I never thought science is the whole story. What knowing mathematics and physics did was lead me to develop my own voice in philosophy. I came to Princeton in the fall of 1953. I had taught for one year before that. I was still a logical empiricist. I had written a Ph.D. with Hans Reichenbach, whom I still revere in many ways, and yet by 1957 I was attacking the verifiability of meaning as a major error, and strongly criticizing positivism, and I think part of what caused the change, paradoxically, was teaching philosophy of science every year. Teaching philosophy of science and thinking about science made me realize more and more that the positivist description was a terrible description of science itself.

YPR: *We also want to talk a little bit about your views on ontology and ethics, especially given your book* Ethics without Ontology *in 2004. You've pronounced an obituary on ontology – the study of what exists – and claimed that it is disastrous to try to base ethics on ontology.*

HP: That attack was an attack on "ontology" in the sense that W.V.O. Quine gave to that word, which is now the standard sense in analytic philosophy. Of course ethics should be based on an account of what we are; I wasn't criticizing that. What I did criticize in *Ethics without Ontology* was two ways of ontologizing ethics. One is to look for mysterious entities. So for example if you talk about human rights, then there must be objects which are rights, and then the question becomes, what is the nature of those objects? If you talk about good and bad, there must be such an object, a property of goodness, and then you ask, is it like yellow, is it not like yellow, in what respects is it like yellow, and the whole schmear is upon us. I regard this search – this effort to base parts of science or human life on objects unknown to either science or common sense – as a disaster.

YPR: *If ontology is not the foundation for ethics, what do you see as the alternative foundation?*

HP: First of all, we have to rethink what ethics is. If we approach ethics from the standpoint of the semantics of ethical terms, I think one can see very clearly that much of the problem for philosophers comes from the assumption that the correct semantics for ethical terms must enable us to show that the evil man is irrational. That seems to me too ambitious a hope. The idea that the right account of ethics must show that the evil man is making some kind of logical mistake – so if you found the right argument, then the evil man would either become a good man or would slink away unhappy because he'd been exposed as a fool – I don't believe there is such a magical argument. Ethics cannot simply be reduced to reason. The idea of showing the evil man to be irrational is the Platonic hope. That hope has to be abandoned once and for all.

I think ethics is something human. I think it rests on human interests, and I don't even think it's universal in the way some had hoped. But that doesn't mean that I'm a cultural relativist. I think one can have a kind of moderate view which acknowledges that ethics is something that changes with time, that it has a history, that it has as a dependence on culture, but does not draw nihilist conclusions from that. At least that would be my program. What I think is that ethics rests on a set of human interests, which have shown themselves to be of very wide appeal. Amartya Sen has a new book, *The Idea of Justice*, which is not yet published, but which he has let me read. He says that while it is true that institutional democracy is often said to be a Western institution, and indeed modern democratic institutions are a Western product of the past two centuries, democracy at the level of local governments and certain democratic aspirations and ideals go back many centuries. King Ashoka of India, who is one of Sen's examples, promulgated the virtue of tolerance, for example. So while democracy in the institutional sense is new, the search for, the aspiration for, democracy is at least as old as civilization.

I would say that Thomas Scanlon's claim – that a lot of suspicion of the objectivity of ethics is actually a suspicion of the objectivity of the reasonable – is correct. And I think a problem with that kind of skepticism is, once one has become skeptical of reasonableness, or the notion of reason, then consistency demands that one be skeptical of all knowledge, including scientific knowledge, because science for sure rests on the idea of reasonableness, which in turn rests on a whole set of values—epistemic values. In my recent work I've talked about the triple entanglement of fact, value, and convention. The entanglement of fact and value at the level of epistemic values is almost transparent. Scientific theories are selected in part on the basis of virtues such as what the great physicist Paul Dirac called "beauty" and Einstein referred to as "the inner perfection of a theory." The positivist idea that science can proceed by just an algorithm called "inductive logic" has come to nothing.

So the strategy I follow is to fight on two fronts: on the one hand, against total skepticism about value. I think total skepticism about value undermines fact as well. It undermines the so-called paradigm facts of science. On the other hand, I agree with Kant when he said that reason loses its claim to its rightful territory if it claims too much. The client is an unruly client. This is the great central metaphor of the *Critique of Pure Reason*. You have to get the client to moderate her claims, and make them reasonable claims. And once she does, at that point, you have to really fight back against the skeptic. But you don't beat the skeptic by trying to defend an exaggerated notion of objectivity, or exaggerated claims. Fighting on both fronts is my strategy.

YPR: *It seems that under the number of claims that could be deemed reasonable, there might be many, and they might potentially conflict. How do we decide between these claims?*

HP: The best method we have – parodying Churchill, you might say "the worst of all possible methods except for those others which have actually been tried" – is well-informed democratic discussion. Obviously if we use that method, we will not end up with a complete ordering of all the relevant preferences. We will end up at best with a partial ordering. Naturally, there will be disagreements even over the partial ordering. But the hope is – and I think it's a well-grounded hope – that there'll be a lot that we can agree on if we can agree to be reasonable. This was of course Rawls's great faith all his life. If we are willing to not rely on prejudice, on preconceptions, on what Michelle Moody Adams calls "affected ignorance," then I think we make progress via reasonable discussion among people who share the basic interests of morality – at least the interest in living together under commonly agreed rules. That's the element that's right in Contractarianism. I'm not a Contractarian. But I think the Contractarians put their finger, in my view, on one central interest of morality.

Generally most ethicists put their finger on one or another central interest of morality. They are generally unwilling to admit that there's more than one, or else they try to derive all the others from whichever they prefer. So if there are a number of central interests of morality – equality is one, justice is another, compassion for the weak is another – then these interests do sometimes conflict. Absolutely. And then the best we can do is to try to discuss what to do in light of these interests. But if someone is not interested in being compassionate at all, or in being just at all, then the best you can do in such a case is to try to show her the appeal of the compassionate way of life or of a way of life informed by justice. And let me say that here Utilitarianism has an element of truth. One respect in which ethics is objective is that it is a fact that abandoning it altogether is generally worse for just about everyone, even in strictly utilitarian terms. As a short and oversimplified answer to the question "in what sense is ethics objective" – Utilitarianism was a good first stab.

YPR: *To what extent should being ethical define our lives? The twentieth-century philosopher Emanuel Levinas talks about how we have infinite responsibility for everyone else. You respond that "to be only ethical is to live a one-sided life." How then can we determine the force and extent of ethical demands, and what other interests should we keep in mind when making our life decisions?*

HP: This is a question that Habermas and I have argued about for years (I didn't know Levinas personally). Jurgen Habermas says that you can't be both a Kantian and an Aristotelian. I say no, Aristotle without Kant is blind, and Kant without Aristotle is empty. Aristotle faces the question you ask from the first chapter of the Nichomachean Ethics: how do you reconcile your legitimate concern with personal fulfillment, with eudaimonia, with flourishing, how do you reconcile that with concern with the *polis*. I believe that this question must receive different answers depending on whether the question is addressed to the whole community or to the individual. For the community, the only answer, as I said, is that democracy is a very fallible way, but it's the best way we have. But for the individual, it is a different matter. Subject to Rawls' first principle – that you claim only those basic liberties that you're willing for everyone to have; that's the Kantian moment in Rawls – subject to that Kantian constraint, the answer is that this is a question of deciding who you are. No one else can do that for you. And Levinas did decide who he was, and I respect him for that, but as a universal prescription, taking on "infinite responsibility" underrates the claims of eudaimonia.

YPR: *Why do you think many people find it harder to accept the objective truth of ethical statements such as "murder is wrong" than to accept the objective truth of mathematical statements such as "2+2=4"?*

HP: Perhaps it would be better to ask why so many people think physics is objective and ethics isn't. Physics too isn't "mathematically certain" but it is often the paradigm

of objectivity. Many people think "There's science over there, and that's clearly safe, that's clearly objective—I don't know anything about it, but it's clearly objective, and anyway the gadgets work." And, by the way, the fact that we need the myth of "postmodernism" shows a lot. What it's based on, I believe, is grumpy pessimism. I mean, the West was on top of the world. It ran the whole show, and it had fantasies of empire and perpetual progress, and so on. Since World War II, it's become more and more clear that we in the West have to share this world with non-Westerners, not to mention with minorities in the West, or majorities, like women. And the West's loss of self-confidence is now being turned into the idea that there was nothing there anyway, progress was a myth, and objectivity is a myth.

YPR: *Let's talk a bit about your interest in religion. You've just written a book called* Jewish Philosophy as a Guide to Life. *How did religion and Judaism in particular come to be such a significant part of your life, both personally and as a philosopher?*

HP: I was an atheist in my childhood and for some time after, perhaps because my father was a communist but he was a communist with a soft streak for religion. My father had this soft spot for religion, and even an attraction for Roman Catholicism, all his life. So while I grew up with Communist parents, I did not grow up with religion-bashing parents. So I think that was part of it, that although my upbringing was an absolutely secular one, I was not a hater of religion in any way. In fact I was drawn to it. In college, my favorite existentialist was Kierkegaard, who seemed incomparably deeper to me than Sartre. Sartre was the fashionable existentialist after World War II, but it was Kierkegaard who really spoke to me. Some religious literature spoke to me, some religious poetry did, and so on. But the involvement with Judaism really began in 1975 when my son told my wife and myself that he wanted to have a bar mitzvah. And then I formed a friendship with Rabbi Gold [Director of Harvard Hillel from 1958 to 1990.] Under Rabbi Gold's guidance, my wife and I fell in love with Judaism, with Jewish culture, Jewish history, and so on. I should mention that my wife is a survivor. She was not in a camp, but she survived Nazi Germany, and when Sam said he wanted a bar mitzvah she said she felt as if she'd spit in Hitler's eye. She said "Hitler wanted to kill me, and now it's going to go on for at least a generation after me."

YPR: *How would you respond to the view that religion and philosophy are irreconcilable, because religion involves some sort of acceptance of a revealed tradition and a commitment to living one's life in pursuit of God's vision for the world, while philosophy requires a life of reasoned questioning that precludes an acceptance of anything based on faith?*

HP: Atheists and fundamentalists share the view that any religious person who isn't fundamentalist isn't really religious. Each casts religion the same way – the atheist so that religion will be vulnerable. I'm sure among the ultra-Orthodox you can find a lot of

that, but most of the religious Jews I know are not fundamentalists, including the Orthodox Jews I know. Fundamentalism seems to be much more common in certain Protestant sects. And I suspect most American Protestants, if you press them on it, would turn out to have a much vaguer sort of religiosity than the fundamentalists do. Then, of course, the critic's next move is: well if religion is vague, it has no content. But that assumes something that Wittgenstein rightly attacks. I very much admire Wittgenstein's lectures on religious belief, by the way; they are almost the best of what he wrote. Wittgenstein really attacks the idea that you can think of religion as a theory, which Kierkegaard whom, as I say, I liked since college, also attacks. Religion isn't a theory. It's a way of life. It's a way of navigating the "stream of life."

Obviously what's important to one person is not important to another. There are people for whom what's important is the literal truth of the claims of their sect. Well, to that I say, as William James did, "be tolerant of anything that's not itself intolerant," but that's not for me. For me religion is a way of life, and the Jewish religion is one particular spiritual way of life. But I do believe in the reality of spiritual experience. Elizabeth Anscombe was a good friend, one I much admired. I didn't share her religious views at all. But she said when I first met her in 1960 that the difference between the religious person and the atheist is like the difference between the person who sees the stained-glass windows from the outside and the person who sees the stained-glass windows from the inside. That's an image I've never forgotten. Maybe for Judaism stained-glass windows aren't the best image, but I see the Jewish religion from the inside, not from outside. It is a vision that illuminates my life and my way of experiencing the world. That's not a theoretical vision. Theoretically I'm a naturalist.

YPR: *You've been involved quite a lot in political activism – sometimes controversial political activism.*

HP: Yes, and sometimes mistaken. Like joining the PLP [Progressive Labor Party].

YPR: *So, what do you see as the proper relationship between philosophy and the practice of politics?*

HP: In Sen's *The Idea of Justice*, which, as I mentioned, is coming out, he discusses this question a good deal. Sen criticizes Rawls, although with great reverence; Rawls put the theory of justice back on the map. Nevertheless Sen criticizes Rawls and Nozick and Nagel for an assumption they share. They assume that if we have the right description of an ideally just state – and one of Rawls' assumptions is that in the ideally just state, all the citizens behave in accordance with the idea of the just state – if we know what an ideally just state and ideal behavior are, that will help us with questions of comparative justice. We will be better placed to answer questions of the form "in this imperfect world, would we be making things more or less just if we did X?" And Sen argues that there's no reason to think that's true: a theory of the ideally just society is neither

necessary nor sufficient for a theory of comparative justice. So part of the answer is that one has to descend from the level of always trying to describe the perfect society. What Sen is arguing is that there is room for theoretical discussion which is more practical in the sense of being addressed to the questions of comparative justice that arrive in the real world – the real world as a whole, the global context. Theories of ideal justice also tend to be theories of ideal justice for a single nation-state. But a realistic theory cannot presuppose a single nation-state, nor can it be content with just the minimal requirements of Rawls's Law of Peoples. In this regard we might respect and read Dewey's moral writings, which I think are tremendously impressive and almost not discussed by moral philosophers nowadays. Dewey is precisely always at that intermediate level. He speaks about ideals, yes, but about how ideals bear on the here and now. One of my favorite essays of Dewey's – it's only about five pages, maybe ten pages – written in 1893, I think, is called "Teaching Ethics in the High School," which I would recommend to every philosophy major.

The other error that philosophers sometimes make, the opposite error from Rawls's, is to think that your philosophical theory tells everyone what to do right now, and it obviously can't fail, because philosophers are always certain. But the problem in philosophy is to recognize that the translation of philosophical ideals into practical policies is something that requires a nose for what Dewey calls "problematic situations," that requires democratic discussion, and that requires fallibilism. As Michelle Moody Adams pointed out in *Fieldwork in Familiar Places*, philosophy can't be any longer the judge, the arbiter, of what all the non-philosophers say; it's not that you can bring what these other people say to the Supreme Court, which is philosophy, and philosophy can tell you who is right. But philosophy can be a distinctive participant. That seems to me exactly the right account.

YPR: *I know you haven't perhaps revised your opinions over the years as much as people suggest, but all the same it does seem that you've been able to keep an open mind about various questions by returning to them and rethinking them. It's often hard for people in philosophy to avoid being wedded to their opinions. Do you have any thoughts on how to keep an open mind, which seems to be such an important part of the philosophical endeavor?*

HP: First of all, I wouldn't deny that I've changed my views on at least one large metaphysical question. I am thinking of my so-called "internal realism," which I defended for about fifteen years…But if anyone thinks that they have the answer to all philosophical questions, they must be in the grip of some picture whose problems they are unable to see, or else they have to be blinded by egotism. The desire to be the Newton of philosophy, which Hume already had – to me that's always been foreign. As soon as I publish something I start thinking 'was I really right'? I keep coming back years later and thinking 'is that really right' or 'there's this difficulty' or 'is the way I met

this difficulty really good,' and so on. For me the philosophical life should be the life of questioning one's own ideas first and foremost. And I think one does see that very much in the work of the great Greek philosophers, Plato and Aristotle. One is aware of that – especially in Plato, but I think also in Aristotle, that they are going over their own arguments and raising difficulties.

YPR: *You've argued that we can make progress in Ethics through what's called a Pragmatist enlightenment, through which we apply "situated intelligence" to political and social problems. Could you talk a little more about what you mean by progress in that context, and on what ethical questions, if any, do you see progress being made now?*

HP: Let's take one great question which has been with us since Marx. In fact it was raised before Marx; Hegel was very important in this regard. This is the issue of socialism and capitalism. I think Hegel's lectures on the philosophy of history are most impressive not for the grand picture—the grand picture is impressive, obviously, and to some extent even though all historians say they're past Hegel, all the historians still critically fall into the trap of trying to say what the great pattern is. It's very hard to resist. But the grand picture is not what impresses me about Hegel.

What impresses me about Hegel is his analysis of the modern nation state. Hegel was the first philosopher to treat the state as something with a history. He analyzed the modern nation-state and saw that it is not like the Greek city-states or the Persian empire or the Roman Empire. The modern state is inseparable, he thought, from capitalism. And Hegel thought that the great function of the modern nation-state is to negotiate class conflict. This is not often mentioned, right? You probably are not told that in your history classes. Yes, Hegel thought class conflict should end in compromise, and he thought that the state and what he taught us to call civil society have roles to play in finding compromises. And the Marxists of course reject that and opt for making the ideal state via revolution. By the way, Dewey said that his disagreement with Marxism was not about class struggle. He said, "I believe in class struggle. The mistake of the Marxists is they think that once you've nationalized something, you've socialized it." This is a remark one can think about for a long time and many times.

It's obvious that late capitalism has enormous problems. It's also obvious that what you might call late socialism has enormous problems. I think the problems with which Hegel and Marx were wrestling are still very much with us, and in a sense the great debate between Rawls and Nozick decades ago, in the 1970s, *A Theory of Justice* versus *Anarchy, State, and Utopia*, was a reflection, an echo at the theoretical level, of this question: does anything come after capitalism? If so, what would it look like? That certainly is a question on which we'd better be able to make progress.

YPR: *If you could ask one question to one figure in the history of philosophy, to whom would you ask a question, and what would it be?*

HP: That's not a natural question for me, because I don't view the truth as lying in the past. There are figures in the history of philosophy I revere. If you asked me "who are the four-star philosophers?" I'd say maybe there are only three of them: Plato, Aristotle, and Kant. And then you have a larger group, probably under a dozen, of three-star philosophers, like Descartes. But I'm not one of the people who get tired of the excesses of present-day philosophy, either present-day analytic philosophy, or present-day postmodernism, and then start touting their philosopher of the past. I mean, I love Dewey, but Dewey was a man of his time. I think we've seen further into these questions. It's true that we've forgotten some of Dewey's insights. But the idea of going back to Dewey, or Peirce, or Alfred North Whitehead, or somebody else's favorite dead philosopher, and saying "Oh, he had the answers," I don't believe that. One of the things I like about Dewey is that he said very explicitly and very early that a philosopher can't be expected to produce infallible answers. It's enough if she produces better answers to the relevant problems of her time. And I think that is the goal we should have.

About the Authors

SAMUEL BAGG graduated from Yale University in 2009 with a B.A. in Ethics, Politics and Economics, and he will be teaching history and philosophy at the Ecole D'Humanité in Goldern, Switzerland next year. Hailing from Durham, North Carolina, his plans for the more distant future are unclear at the moment, but for now, he enjoys teaching, talking, writing, thinking, and revisiting the keys of his erstwhile master, the piano. He doesn't like giving shout-outs, but quietly acknowledges his debt to the people who have helped him find comfort in his newfound confidence.

FAHD HUSAIN, born and raised in Karachi, lives and works in Montreal where he divides his time between contemporary philosophy and fine-art photography. Having recently graduated from McGill University, he is currently looking for intriguing and innovative ways to apply his philosophy degree while he debates the benefits and possible drawbacks of further study. His academic interests lie somewhere between the shifting nexuses of "postcolonial" theory, twentieth century French philosophy and liberal political thought. He can be reached at paintednegative@gmail.com.

NAL KALCHBRENNER is a student of philosophy and symbolic systems at Stanford University. His interests involve logic, the foundations of mathematics and of brain and behavior sciences. He wishes to thank Michael Friedman for his vivid lectures and insightful suggestions.

BENJAMIN HERSH studies the foundations of psychology and the history of science at Stanford University. He plans to work at a Bolivian Zoo in the near future, though he will likely walk an academic path further down the line. When not arguing, reading, or pacing frantically in a hallway, he may be found writing poetry on academic whiteboards, charting maps, or painting portraits on found objects. He is deeply indebted to Dagfinn Follesdal for teaching him a more genuine way to philosophize.

PUNEET DHALIWAL is an undergraduate student of Philosophy, Politics, and Economics at the University of Warwick. His studies encompass modern and contemporary political theory, ethics, social choice theory, and economics. His academic interests lie within political theory, political philosophy, and moral philosophy and he plans to pursue postgraduate work in political theory. His wider interests include journalism and political activism. He would like to thank Kai Spiekermann for his initial encouragement, and Mark Carrigan, Chris Rossdale, and Beth Smith for their helpful comments.

SPECIAL THANKS

Michael Della Rocca, Shelly Kagan, and the Yale Philosophy Department
Ian Shapiro and the Yale Program in Ethics, Politics, & Economics
Yale Undergraduate Organizations Funding Committee

Support of this publication was assisted by a grant from the John K. Castle fund. The Castle Lecture and Publication fund was endowed by Mr. John K. Castle. They honor his ancestor the Reverend James Pierpont, one of Yale's original founders.

This issue of THE YALE PHILOSOPHY REVIEW was printed by Tyco of New Haven, Connecticut. The typeface is Garamond.