The Yale PHILOSOPHY REVIEW

An Undergraduate Publication

ISSUE 1 | 2005

The Yale PHILOSOPHY REVIEW

An Undergraduate Publication

Issue I, 2005

Email the editors at: editors@yalephilosophy.fastmail.fm
For our online edition and submission guidelines, visit <www.yale.edu/ypr>.

EDITORS' NOTE

"Unless," I said, "the philosophers rule as kings or those now called kings and chiefs genuinely and adequately philosophize... there is no rest from ills for the cities... nor I think for human kind..."

Plato, *Republic*, Book V

Most today find little comfort in Socrates's solution to humanity's problems: put the philosophers in charge. As people concerned solely with Truth, philosophers tend not to have much in the way of practical know-how, the common complaint goes. We need only recall the Greek philosopher Thales—said to have fallen down a well while gazing at the stars—to realize how far back the image of the Absent-Minded Philosopher stretches.

As students of philosophy, we stand in the threshold between pragmatism and *Sophia*. From this vantage, we can appreciate both views of Thales: as stumbling oaf and as searcher of Truth. Thus, we are simultaneously compelled to pursue truth, and to make this pursuit meaningful to those who doubt its importance. To this end, we have founded *The Yale Philosophy Review*. You hold our first issue. It represents the dedication of students and faculty, but also a tradition of discussion and dialectic. The issues discussed—from Kantian ethics to the nature of "things"—are at the center of the Western philosophic project, but also have much to tell about our lives and our world.

In that vein, we dedicate this first issue to the philosopher-kings—past, present, and future...

X.A. Botero, J. Martin, & A. Srinivasan
Editors-in-Chief

FOREWORD

It is a great pleasure to see the publication of the inaugural edition of *The Yale Philosophy Review*. If the number of submissions is any indication, the *Review* has already generated considerable interest. This should not be at all surprising: the philosophy major at Yale, as at so many other institutions, is thriving and the quality of undergraduate work in philosophy—and, no doubt, the quality of undergraduate teaching!—is often nothing short of inspiring. As I and many other philosophers can attest, so much of the work of philosophers is developed and honed through lively philosophical give-and-take with one's students. Indeed, I would venture to say that not a single philosophical idea of substance that I have had has been unaffected by the challenges I have encountered through the work and thought of my undergraduate students. The support that the Yale Philosophy Department gives to this journal is but a small token of the gratitude we feel toward our students. The excellent and varied papers in this inaugural edition will be—I hope and expect—an excellent way to continue the dialogue that is essential to philosophy and the editors are to be commended for their judgment and for the hard work and devotion they have poured into this project.

Michael Della Rocca
Chair, Philosophy Department
Yale University

THE Yale Philosophy Review
Issue 1, Spring 2005

On Being Able to Know Contingent Moral Truths

The Divine Command Ethics of John Duns Scotus

Sydney Penner
Yale University

From Plato's dialogue *Euthyphro*, we have a famous and troubling problem: is the good indeed good because it is loved by the gods, or is it loved by the gods precisely because it is good? Traditionally, most philosophers have responded to Euthyphro's dilemma by affirming that the good is loved by God because it is good. In contrast, John Duns Scotus, the 13th Century theologian and philosopher, is often interpreted as a voluntarist who defends the opposite claim: that the good is good because it is loved by God. In this paper, Penner argues that recent Scotus interpreters, such as Thomas Williams, who present the philosopher as a strong voluntarist, are inattentive to a number of key passages by Scotus that appear to be inconsistent with such a stance. If moral dictates are not simply contingent on God's will, this paper argues, it becomes easier to see how Scotus can claim that moral goodness is accessible to natural reason. Finally, Penner argues that strong voluntarism is susceptible to a serious epistemological problem: it makes morality contingent upon God's willing, but leaves reason bereft of the knowledge of which god's commands to obey.

Since antiquity at least, theistic philosophers have been bothered by the famous dilemma Socrates posed to Euthyphro: "Is the pious being loved by the gods because it is pious, or is it pious because it is being loved by the gods?"[1] Plato answered the dilemma by saying that the gods love the

[1] Plato, *Euthyphro*, 10a.

pious because it is pious. The good is an eternal form that is not contingent on the gods willing it to be one way or another. Plato's answer remained dominant for a millennium and a half, but with John Duns Scotus's response to Thomistic Aristotelianism during the scholastic period, the other answer to the dilemma returned to prominence.[2]

In this paper I will discuss the voluntaristic ethics of Scotus, specifically in light of an epistemological worry about strong versions of voluntarism. The epistemological worry is this: how can we know what the good is if the good is somehow dependent on what God's will is? Clearly, if divine command ethics is simply defined as any ethics in which moral obligation is dependent on God's commanding, then there is substantial latitude left for working out the picture of the relationship between the content and obligatoriness of morality and God. Different pictures will vary in their susceptibility to epistemological worries. An account in which moral goodness is generally seen as intrinsic—i.e., logically prior to God willing it to be good or bad—and in which it is only obligation that derives from God's commanding,[3] will be less susceptible to these worries than an account in which all the content and obligatoriness of morality is dependent on God's commanding. In an extreme version of the latter, I fail to see how any proposal, including appeals to divine revelation, could adequately answer the question of how we can know, as it does not even provide any reason for us to think that God is any better or worse if he is systematically misleading when revealing the content of morality to us.

There are several reasons why the ethics of John Duns Scotus are an

[2]See Günter Stratenwerth, *Die Naturrechtslehre des Johannes Duns Scotus* (Göttingen: Vandenhoeck & Ruprecht, 1951), 5.

[3]For an example of such a reading, see Frederick Copleston: "Inasmuch as the divine intellect, considered as preceding an act of the divine will, perceives the acts which are in conformity with human nature, the eternal and immutable moral law is constituted in regard to its content; but it acquires obligatory force only through the free choice of the divine will. One can say, then, that it is not the content of the moral law which is due to the divine will, but the obligation of the moral law, its morally binding force" (*The History of Philosophy*, vol. 2, *Mediaeval Philosophy: Augustine to Scotus* [London: Burns Oates & Washbourne, 1950], 547).

especially apt choice for this discussion. First of all, Scotus provided a profound picture of what a divine command ethics might look like. Additionally, the epistemological worry is easily raised when studying his account, as he appears to leave much of morality contingent on God's will, and yet insists that moral goodness is accessible to right reason. This dual claim cannot fail to raise the epistemological question. Furthermore, in the secondary literature on Scotus, there is vigorous debate about how to understand the voluntarism in Scotist ethics—a debate that is directly linked to the epistemological worry.

Interpreting Duns Scotus

The difficulty with interpreting Scotus on morality is the presence within his work of two lines of thought which are not easily reconciled. He clearly states that moral goodness is accessible to natural reason, that it possesses a *convenientia* or fittingness that reason can recognize. This poses no problem for the first two commandments of the Decalogue, as Scotus thinks that the first practical principle, *Deus diligendus est*, is evident from its terms. Moral principles evident from an understanding of their terms or following deductively from such evident principles are part of natural law strictly speaking; thus, there is little difficulty in seeing how they are accessible to natural reason. But Scotus does not think that the commandments of the second table of the Decalogue are either evident from their terms or deducible from evident principles. He appears to hold that they are contingent on God's commanding them and that God could freely have commanded otherwise. This is where it becomes difficult to see how moral goodness is accessible to natural reason. If in identical circumstances God can make killing licit or illicit, how can natural reason recognize which is in fact the case? It is the two lines of thought leading to this difficulty that I want to examine more closely in this paper, although no attempt will be made to present a comprehensive account of Scotus's nuanced ethical thought.

One of the primary texts on the nature of moral goodness is *Ordinatio* I dist. 17, nn. 62-67 in which Scotus compares the quality of moral goodness

to the quality of beauty. It is worth quoting at some length:

> One could say that just as beauty is not some absolute quality in a beautiful body, but a combination of all that is in harmony with such a body (such as size, figure, and color), and a combination of all aspects (that pertain to all that is agreeable to such a body and are in harmony with one another), so the moral goodness of an act is a kind of decor it has, including a combination of due proportion to all to which it should be proportioned (such as potency, the object, the end, the time, the place, and the manner), and this especially as right reason dictates should pertain to the act, so that we could say of all these things that it is their conformity to right reason that is essential... For whatever be the object, and howsoever the act has to do with it, if it is not performed according to right reason (for instance, if right reason did not actually dictate its performance), then the act is not good.[4]

A number of observations might be made about this text. First of all, it reveals Scotus's emphasis on the agent willing to do the morally good act. Merely happening to perform an act with good consequences is not sufficient. The agent must use her reason to judge the goodness of an act and then intend to perform that act with the right end in mind.[5] Indeed, in *Quodlibet*, q. 18, Scotus affirms Pseudo-Dionysius's declaration that "Good

[4] *Ord.* I, dist. 17, n. 62: "*Dici potest quod sicut pulchritudo non est aliqua qualitas absoluta in corpore pulchro, sed est aggregatio omnium convenientium tali corpori (puta magnitudinis, figurae et coloris), et aggregatio etiam omnium respectuum (qui sunt istorum ad corpus et ad se invicem), ita bonitas moralis actus est quasi quidam décor illius actus, includens aggregationem debitae proportionis ad omnia ad quae habet proportionari (puta potentiam, ad obiectum, ad finem, ad tempus, ad locum et ad modum), et hoc specialiter ut ista dictantur a ratione recta debere convenire actui... quia quantumcumque actus sit circa obiectum qualecumque, si non sit secundum rationem rectam in operante (puta si ille non habeat rationem rectam in operando), actus non est bonus.*" When included in his book, I will also cite the page number of Allan B. Wolter, O. F. M., *Duns Scotus on the Will and Morality* (Washington DC: Catholic University of America Press). Here AW 206-207.

[5] Cf. Kant's emphasis on the good will in *Grundlegung zur Metaphysik der Sitten*.

requires that everything about the act be right, whereas evil stems from a single defect," though Scotus does go on to point out that the absence of any one of the required circumstances only results in privative badness. That is, for an act to be contrarily bad there must be some positively bad circumstance. For example, if one gives alms to the poor but lacks a good end such as to love God or to help one's neighbor, the act is privatively bad. If, however, one gives alms with a positively bad end such as vainglory, the act is contrarily bad.[6]

The relevant point is that moral goodness is not a simple absolute property possessed by acts in the manner that objects, for example, possess properties like shape or color. Rather, it is a complex relational property, like beauty, that can be recognized as a *convenientia* between a variety of other properties. While it is clear that Scotus thinks that moral goodness is accessible to right reason (otherwise it would be futile to suggest that right reason should dictate acts), it is still unclear exactly what this accessibility means. Specifically, does moral goodness necessarily follow if all the circumstances are right? After all, certain relational properties must follow from certain sets of conditions. For example, "being between two other objects" is a relational property and it is a property that must be possessed by the middle object in a row of three objects. If the *convenientia* that results in beauty or moral goodness follows with this sort of necessity, then it is easy to see how it would be in principle accessible to right reason.

The problem with this is that it seems to make moral goodness prior to any act of judgment on God's part. That is, it is difficult to see how a given act which features all the right circumstances for a complete *convenientia* could be proscribed by God as illicit except as illicit by positive law; more obviously, it is difficult to see how an act that failed to meet the requirements of *convenientia* could be made licit by God. Yet in other texts Scotus does want to say that God has this sort of freedom to make acts licit or illicit. But if God's act of commanding is necessary as well, it seems that the list of circumstances cited in *Ord*. I, dist. 17, n. 62, is insufficient. A possible way out of this difficulty is if Scotus is read as including God's act of command-

[6] *Quod*. 18. AW 210-219.

ing as part of the domain of circumstances that establishes *convenientia*. It would be striking, however, that Scotus makes no mention at all of God in this passage if that were to be a critically decisive circumstance.

The analogy with beauty is tantalizing, but unfortunately Scotus does not develop a detailed aesthetics anywhere and so not too much can be drawn from the analogy. Francis J. Kovach has done a preliminary survey and analysis of references to beauty in Scotus's writings and concludes, unsurprisingly, that Scotus accepted the standard medieval assumption that beauty is an objective empirical fact.[7] This still leaves open the question of what exactly makes an incident of beauty, or moral goodness, an objective fact rather than a subjective response. One obvious option is that beauty and moral goodness are objective, albeit relational, properties inhering in objects or acts. Again, this appears to leave moral goodness as prior to God's free act of judgment. The other option is to say that beauty and moral goodness are objective from our human perspective, insofar as certain judgments have been made by God and it is our responsibility to conform to God's judgments. With this option, God was free to judge differently but the divine judgment creates an objective standard.

Two key texts that are often cited to emphasize Scotus's voluntarism[8] are his discussion of natural law in addressing a question about the obliga-

[7] "Divine and Human Beauty in Duns Scotus Philosophy and Theology," *Studia Scholastico Scotistica* 5 (1972):445-459.

[8] I follow Thomas Williams's definition of "voluntarism" when I use the term: "(i) the goodness of almost all things, as well as the rightness of almost all acts, depends wholly on the divine will and (ii) what God wills with respect to those things and those acts is not in turn to be explained by reference to the divine intellect, human nature, or anything else" ("The Unmitigated Scotus," *Archiv für Geschichte der Philosophie*, 80 [1998]: 162-181). I will take the "depends wholly on the divine will" loosely enough to be able to consider more "mitigated" interpretations of Scotus such as John Hare's in which moral properties are seen to supervene on other properties as voluntarist. I see the key feature of voluntarism to be the idea that in identical circumstances God could make a given action either licit or illicit without the dictate necessarily becoming merely positive law thereby. It is important that this definition of voluntarism be kept in mind since there are certainly authors who would use the term "voluntarism" for their reading of Scotus but would reject the sense just described.

toriness of confessing sins to a priest (*Ord*. IV, dist. 17) and his discussion of whether all the commandments of the Decalogue belong to natural law (*Ord*. III, suppl., dist. 37). In these texts Scotus delineates explicitly what he sees natural law to be, setting his own view against the Thomistic conception. Concerning the definition of natural law, Scotus says:

> Now, a practical truth of natural law is either one whose truth value can be ascertained from its terms (in which case it is a principle of natural law, even as in theoretical matters a principle is known from its terms) or else one that follows from the knowledge of such truths (in which case it is a demonstrated conclusion in the practical order). And strictly speaking, nothing pertains to the law of nature except a principle or a conclusion demonstrated in this fashion.[9]

The first practical principle, *Deus diligendus est*, is the practical truth that Scotus thinks can be ascertained from its terms. That is, a proper understanding of what the term "God" means will show that God is the *summum bonum*. Since it is evident that the good ought to be loved, it is clear that God as the highest good ought necessarily always to be loved.[10] No contingent premises need to be drawn in to see that this is so. Furthermore, principles that follow necessarily from such first practical principles also belong to natural law strictly speaking. It is easy to see why they would be given the same status since they also do not need to draw on any contingent premises and hence carry the same necessary force. That they must follow necessarily in order to qualify as strict natural law is evident from Scotus's description of natural law in *Ord*. III, suppl., dist. 37: "What pertains to the law of nature is either a practical principle known immediately from

[9] *Ord*. IV, dist. 17: *"Illud autem est verum practicum de iure naturale cuius veritas est nota ex terminis-et tunc est principium in lege naturae, sicut et in speculabilibus principium notum ex terminis-sive quod sequitur evidenter ex tali vero sic noto, cuiusmodi est conclusio practica demonstrata. Et stricte loquendo nihil aliud est de lege naturae nisi principium vel consilio demonstrata sic*." AW 262-263.

[10] More precisely, it is the precept "Do not hate God" that belongs to the natural law strictly speaking. See *Ord*. III, suppl., dist. 37 (AW 282-83).

its terms or necessary conclusions that follow from such principles."[11] The reason for their necessary truth is that they have the ultimate end, God, as their immediate object.[12] Given that these practical truths of natural law are necessary truths, God cannot make them false. This is a crucial point: God cannot make what these principles say is good to be anything but good and cannot make what they say is illicit licit.[13] There can be no dispensations from natural law strictly speaking.

But according to Scripture, or at least Scotus's interpretation of Scripture, God did indeed provide dispensations from numerous commandments, including commandments of the Decalogue. The most frequently discussed examples include the commands for Abraham to sacrifice his son Isaac, for the Israelites to despoil the Egyptians, and for Hosea to marry an adulterous wife. Clearly, if a practical truth of the natural law has necessary force, there can be no dispensation from it; unsurprisingly, these apparent dispensations were a widely discussed problem among medieval philosophers, such as Thomas Aquinas, who considered all the precepts of the Decalogue to be part of natural law. If they were simply instances of positive law, there would be no difficulty in saying that the lawgiver could dispense from them. But Scotus is also not willing to relegate the majority of the Decalogue's commandments to the status of positive law. Scotus does think that the first two commandments follow necessarily from the first practical principle and so are unproblematically part of natural law strictly speaking. The third commandment regarding the keeping of the Sabbath is somewhat more complicated, for Scotus is unsure whether natural law dictates it be kept on the seventh day in particular. That leaves at least seven commandments—which Scotus refers to as the second table—that are not part of natural law strictly speaking.

This leads Scotus to make his well-known and crucial distinction between natural law strictly speaking and natural law in an extended sense.

[11] *Ord*. III, suppl., dist. 37: "*Quae sunt de lege naturae vel sunt principia practica nota ex terminis vel sunt conclusiones necessariae sequentes ex talibus principiis.*" AW 270-271.

[12] Ibid. AW 277.

[13] Ibid. AW 277.

Not willing to class the second table of the Decalogue as positive law, Scotus has to propose a third category: "[S]omething is said in an extended sense to pertain to the law of nature if it is a practical truth that is immediately recognized by all to be in accord with such a law."[14] What distinguishes these principles from those of natural law strictly speaking is that they are not necessary conclusions. It would be possible for them not to be commanded while still leaving the ultimate end of *condiligentia Deo* attainable.[15] Of course, since they are not necessary, God is free to provide dispensations from them. It is not entirely clear whether there are any circumstantial restrictions on God's freedom to provide these dispensations, but I will examine that more closely later. The other key thing to note here is that the natural law, even in an extended sense, appears to be accessible to right reason. Allan B. Wolter infers from the statement that they are "immediately recognized by all to be in accord with [natural law strictly speaking]," that "they represent conclusions that any reasoning person can readily know or infer" and that "[n]o special revelation…seems to be required."[16] Wolter's conclusion, however, is too hasty. There may be other texts that lead to this conclusion but it does not necessarily follow from this passage. Note that Scotus only says that it is immediately recognizable that the principles of natural law in an extended sense are in accord with natural law strictly speaking. This does entail that any reasoning person should be able to recognize principles *not* in accord with natural law strictly speaking as not being part of natural law in an extended sense either. But the positive counterpart does not follow, since it is in principle possible that more than one mutually incompatible practical principle be in accord with strict natural law. It is not at all clear how reason, in such a case, would be able to judge which principle was in fact part of the actual extended natural law.

A closer examination of the difference between natural law in its strict

[14] *Ord.* IV, dist. 17: "*[E]xtendendo quandoque illud dicetur esse de lege naturae quod est verum practicum consonum principiis et conclusionibus legis naturae in tantum quod statim notum est omnibus illud convenire tali legi.*" AW 262-263.

[15] *Ord.* III, suppl., dist. 37. AW 276-277.

[16] AW 58.

and extended senses might be illuminating. The key distinction is that natural law in its strict sense is necessary and can admit no exceptions. Natural law in its extended sense, on the other hand, is contingent and does allow for dispensations. But what exactly does it mean to say that a practical principle is contingent? There are at least two different ways that could be parsed. A principle could be contingent either on God's free act of commanding or on the contingency of created reality itself (though the latter is, of course, ultimately itself contingent because it depends on a free act of God). Suppose that *A* is a contingent state of affairs that has been actualized and that the hypothetical imperative "if *A*, then *B*" is true, where *B* is an obligation to φ. If the hypothetical is necessarily true, then *B* is contingently true. If, however, the hypothetical itself is contingent, *B* would not follow even contingently. Only the disjunct "$B \vee \sim B$" would follow, which of course would be true even without the premises used. To remove the disjunct we would need further premises such as "necessarily, if *C*, then if *A*, then *B*," where *C* is a command of God contingent on God's freedom, whether or not *A* obtains. Adding this premise to the last argument, as well as the premise that *C* is in fact the case, would allow us to derive that *B* is contingently true—the same conclusion reached in the first argument. So in both of these arguments the conclusion is a contingent obligation to φ.

The first argument is of the same form as saying that a moral principle is contingent as a result of the contingency of created reality generally. The second argument is of the same form as saying that a moral principle is contingent because God is free to choose from two or more principles even in identical circumstances. Though the conclusions may be the same, the distinction between the two arguments is important because of their very different implications. In the first argument, the conclusion follows from a given state of affairs without appeal to any dictate of God. God could, of course, have initially created a different state of affairs, which might then have resulted in different moral principles. Any number of fanciful possibilities could be concocted. If human bodies functioned such that the only way we woke up in the morning were if we had been killed the night before, the general proscription against killing would presumably not be

part of natural law. But once a contingent state of affairs is actualized, the moral principle follows. Thus, in this sense, this argument would represent God loving the good because it is good, i.e., God could not command us to not-φ if *A* were true. This, of course, would be unacceptable to a true voluntarist. The second argument represents a genuine voluntarist position, insofar as the command to φ is contingent on God's free act of commanding, whether or not *A* is true. In this case, then, the morally good would be good because God loves it.

This distinction is important and yet it frequently appears to be confused in the secondary literature on Scotus. All too often, authors who are presenting thought experiments to defend the voluntarist position more plausibly will construct examples in which the circumstances change along with the moral principle.[17] These examples may be interesting insofar as they show us that specific moral laws are not eternally true independent of the contingent circumstances on which they supervene, but these examples fail to engage with the strict voluntarist question at hand. If moral commands are contingent on divine willing, which is not to be explained in terms of human nature or any other part of created reality, then there is no point in constructing elaborate possible worlds to make other moral commands plausible. It must be plausible for there to be other moral commands incompatible with the ones we actually have in a world precisely like our own. It might be suggested that the reason other moral principles might not seem plausible

[17] I use the word "circumstances" as a very broad term so that it refers not just to features of an act like intention, manner, and so forth but also features of the way things are. That is, the fact that killing someone harms them is part of the circumstances. There is some question as to whether Scotus used the term in the same way. It seems to me that this is the most natural interpretation of the "*circumstantiae*" in *Ord*. III, suppl., dist. 37. But it may be that Scotus uses it in a more limited sense to refer to the particulars of a situation such as intention, manner, time, place, and so on. A difference in usage would affect the interpretation of Scotus (though not my epistemological worries about strong voluntarism). It seems to me that the case for Scotus being a strong voluntarist would be weakened if he does in fact always use the term "circumstances" in a more limited sense since it seems to me that a strong voluntarism must affirm that God can make an action licit or illicit even if *all* circumstances, in my broad sense, are the same. I doubt that anyone would deny that the content of morality could alter if the way things are was relevantly different.

to us is because God has given us certain moral intuitions along with commanding a corresponding set of moral principles. This might be a plausible scenario, but it is not clear to me that these intuitions would not be a relevant part of that contingent state of affairs taken into account by the non-voluntarist position.

There are numerous passages from Scotus relevant to determining which sort of contingency he has in mind. Unfortunately, they do not seem entirely consistent with each other even though he frames the question with admirable precision in *Ord.* III, suppl., dist. 37:

> My question then is this. Granted that all the circumstances are the same in regard to this act of killing a man except the circumstances of its being prohibited in one case and not another, could God cause that act which is circumstantially the same, but performed by different individuals, to be prohibited and illicit in one case and not prohibited but licit in the other? If so, then he can dispense unconditionally... But if God cannot cause this act [of killing], which under such and such circumstances was formerly prohibited, to be no longer prohibited, even under the same circumstances, then he cannot make killing licit—but that he did so is clear in the case of Abraham and in many other instances.[18]

This is as clear a statement as could be hoped for. Scotus lays out precisely the sort of scenario I had in mind with my second formal argument and then explicitly affirms it. Were one to make the judgment solely on the basis of this passage, Scotus would clearly come out as a voluntarist. There

[18] *Ord.* III, suppl., dist. 37: "*Quaero ergo an stantibus omnibus circumstantiis eisdem in isto actu 'occidere hominem,' ista circumstantia sola variata per prohibitum et non prohibitum, possit Deus facere quod iste actus, qui cum eisdem circumstantiis aliis aliquando est prohibitae et illicitus, alias esset non prohibitus sed licitus? Si sic, simpliciter potest dispensare... Tamen non potest facere Deus de isto actu, qui cum talibus circumstantiis erat prohibitus, quod manentibus eisdem circumstantiis prioribus prohibitione non sit prohibitus. Ergo non potest facere quod occidere non sit prohibitum-cuius oppositum manifeste patet de Abraham et de multi aliis.*" AW 272-275.

are two problems. First of all, the biblical exegesis concerning the case of Abraham is far from convincing. A prevalent traditional interpretation based on Hebrews 11:17-19 says that Abraham thought that God would raise Isaac from the dead. Surely this would be a significant change in the circumstances!

Of course, this exegetical issue does not have any bearing on Scotus's view. The second problem, though, does. There are numerous other texts by Scotus that suggest a different view than the one just quoted. For example, Scotus's discussion of marriage and bigamy seems to be working on different assumptions. After a discussion of what commutative justice requires, he concludes that "so far as strict justice in the state of fallen nature goes, this contract, if both ends are considered, demands a one-to-one exchange of bodies."[19] It might be noted that this discussion of commutative justice implies that natural reason can, in this matter, arrive at the right conclusion. But Scotus then goes on to say that bigamy could be permitted by God if circumstances were different then they currently are: "For God could have either clarified his law concerning this exchange [i.e., the matrimonial contract] or revoked it in a particular case, and he could have done so reasonably where a greater good would result from its revocation than from its observance."[20] The text here suggests that Scotus thought that changes in the circumstances are what permit God to issue dispensations. Granted, Scotus does not assert this explicitly. Still, the repeated explanations of what changes in the circumstances resulted in God granting dispensations does seem to suggest that if the circumstances had not changed, God could not have granted the dispensations. This stands in contrast to the earlier quotation about God being able to make killing licit or illicit in precisely the same circumstances. Indeed, Scotus goes on to say that in earlier times when it was necessary that humans multiply rapidly, God reasonably gave

[19] *Ord.* IV, dist. 33, q. 1: "*[I]deo de stricta iustitia pro statu naturae lapsae, considerando istum contractum ut est ad utremque finem, requiritur commutatio unici corporis pro unico.*" AW 290-291.

[20] Ibid: "*Potuit enim Deus vel legem suam de commutatione ista vel declare vel in casu aliquo revocare; et rationabiliter in isto casu quando maius bonum provenit ex revocatione quam ex observatione.*" AW 290-291.

the patriarchs a dispensation "so that one man might share his body with those of several women to increase the number of those who worshiped God, something which would not have occurred otherwise."[21] It is significant that this dispensation is in no way arbitrary; rather, there are differences in the circumstances, readily accessible to us, which provide reasons for the dispensation. When Scotus goes on to say why bigamy is not licit in our day, he refers to Christ's statement in Matthew 19 about two becoming one flesh. But he immediately modifies that: "Nevertheless, the reason it is not licit, speaking of justice on the part of the contract and contracting parties, stems from the fact that the principal end does not require it as present. . . "[22] There is a reason grounded in our circumstances that informs Christ's dictate. And if these circumstances changed—for example, if many men died of war or pestilence, while many women remained—bigamy would again become reasonable. As Scotus explains: "And in such a case according to right reason a woman should be willing that her man be joined to another woman that childbearing may occur. All that would be wanting for complete justice would be divine approbation, which perhaps would then occur and be revealed in a special way to the Church."[23] Divine approbation is necessary for complete justice, but again, the dispensation would result from changing circumstances that would provide reasons accessible to us. Dispensations do not occur arbitrarily. Rather:

> [F]or the secondary type [i.e., natural law in the extended sense], dispensation occurs in a situation where the opposite seems to be generally more in harmony with the primary law of nature. And it is just in this secondary sense of natural that monogamy pertains to the law of nature and bigamy is opposed to such. . .

[21] Ibid: "*. . . ut commutaret unus vir corpus suum pro pluribus corporibus mulierum multiplicatione cultores Dei, quae sine hoc non fieret.*" AW 292-293.

[22] Ibid: "*[T]amen loquendo de iustitia ex parte contrahentium et contractus non est modo licitum, quia non est ille finis principalis nunc necessarius. . .*" AW 294-295.

[23] Ibid: "*[E]t deberet tunc velle mulier secundum rectam rationem hoc ut accidat bonum prolis per connexionem viri sui cum alia. Nec deficeret ibi nisi tantummodo completio iustitiae, quae est ex approbatione divina, quae fore tunc fieret, et Ecclesiae specialiter revelaretur.*" AW 294-295.

> But from this it does not follow that in a special case the opposite could not be licit, or even in some cases necessary.[24]

Dispensations from natural law in the extended sense can occur whenever circumstances are such that a dispensation would result in greater good, i.e., would result in greater harmony with the primary law of nature which is directed at our primary good. Obviously there can be no dispensations from natural law in the strict sense, since there can never be changes in our circumstances such that a greater end would be served by dispensing from the natural law in its strict sense. In no possible circumstances would hating God result in greater good. There can be dispensations from natural law in the secondary sense, but the discussion of marriage and bigamy suggests that such dispensations can only occur with relevant changes in circumstances such that there are good reasons for the dispensations. This reading is also suggested by Scotus's discussion of private property as an explication of the difference between the two senses of natural law. There, he argues that the obligation to keep possessions distinct does not follow deductively from the principle that life in community ought to be peaceful. But the obligation to have private property is not arbitrary. Rather, it follows after a contingent premise is drawn: namely, that human nature, in our present sinful condition at least, is such that people having their own possessions is more conducive to peaceful living.[25]

Furthermore, Scotus makes claims about the universality of all natural law that seems to conflict with an understanding of extended natural law as being contingent on God's willing rather than circumstantial features. As Stratenwerth notes: "All norms, Duns Scotus says, which belong to natural law, are the same for all people and in every state. This unchangeableness of the natural right extends itself also to secondary natural law [i.e., natural

[24] Ibid: *"Circa secundum in casu cadit dispensatio,in quo oppositum videtur communiter consonum legi naturae. Et praecise hoc secundo modo monogamia est de lege naturae, et bigamia contra eam. . . Nec tamen ex hoc sequitur quod in casu non possit oppositum esse licitum, immo in casu est necessarium. . . "* AW 294-297.

[25] *Ord.* III, suppl., dist. 37. AW 280-281.

law in the extended sense]…"[26] This is significant because it is not at all clear why natural law in the extended sense would be universally true if it were contingent on God's willing. Why could God not will one set of commands for one culture or for one temporal period, and another set for another culture or period, if the laws are not necessarily true eternal verities?[27] Also, one of the motivations behind forming the category of natural law in an extended sense was to allow for dispensations from it. But if the same natural law applies to all people in all conditions, where are the dispensations? Yet Scotus affirms this universality repeatedly. In a discussion of ceremonial law he contrasts it with natural law by saying: "[W]hatever pertains to the law of nature, either properly or extensively, is uniform…"[28] Elsewhere, responding to Gratian's identification of all divine law with natural law, he says: "Now sacraments, and many other things, are of the positive divine law, and not of the law of nature, for they were not in all states."[29] The reason that some of the divine law is not part of the natural law is that some of it was not binding in all states. In the same discussion, after having given his standard definition of natural law in its strict and extended senses, Scotus goes on to contrast natural law with positive law: "Those principles, which pertain to the law of nature, are the same for everyone, and in all states…"[30] Again, the feature which

[26] Stratenwerth, 107: "*Jene Normen, sagt Duns Scotus, welche zum Naturgesetz gehören, sind dieselben bei allen Menschen und in jedem Status. Diese Unveränderlichkeit des Naturrechts erstreckt sich auch auf das sekundäre Naturgesetz…*" My translation.

[27] A subject that might be worth exploring at some point is how a voluntarist divine command ethics could allow for a certain amount of ethical relativism. That is, it seems that for a voluntarist God might well be free to command different sets of ethical principles to different cultures. This might, of course, be seen as either a desirable or undesirable implication of voluntarism.

[28] *Ord*. IV, dist. 17: "*[Q]uia illa quae sunt de lege naturae sive proprie sive extensive semper sunt uniformia…*" AW 264-265.

[29] *Reportata Parisiensia* IV, dist. 17, q. 1, n. 3: "*[N]am sacramenta, et multa alia, sunt de Jure positivo divino, et non sunt de Jure naturæ, non enim fuerunt in omni statu.*" Quoted from the 1891-1895 printing of the Wadding edition, v. 24, p. 275. English translations of passages from the Wadding edition are my own.

[30] *Rep*. IV, dist. 17, q. 1, n. 4: "*Illa autem, quæ pertinent ad legem naturæ, sunt eædem*

distinguishes natural law in both its strict and extended senses from positive law is its applicability to everyone in all conditions.[31]

Positive law may have an important role to play in coming to an understanding of what, precisely, Scotus means by natural law in an extended sense. After all, Scotus wants three categories: natural law strictly speaking, natural law in an extended sense, and positive law. The meaningfulness of these three categories rests on their distinctness. For a voluntarist, I suspect the difficulty will be to keep natural law in the extended sense distinct from positive law. One of the key features of natural law in the extended sense is its consonance with natural law strictly speaking. But does positive law not also have to be consonant? That is, presumably, if a command of positive law were in opposition to natural law strictly speaking, the positive law would need to be dismissed. So, consonance with natural law strictly speaking is not sufficient to distinguish natural law in an extended sense from positive law. One of the features of positive law, however, is that incompatible commands could all be justly made (though not at the same time). Many ceremonial practices are of this sort, as is made evident, for example, by the fact that different practices were commanded to the Israelites than to the Christians via the New Testament. Perhaps this is not true of natural law in its extended sense. Scotus appears to make such a distinction in discussing whether it is necessary to confess one's sins to a priest:

> But others are beyond those—which are inferred demonstratively from them [i.e., practical principles known from their terms by natural light and all those which follow immediately from them]—which are very consonant with those practical

apud omnes, et in omni statu. . . " Wadding v. 24, p. 276.

[31]The "*statu*" or "conditions" referred to in Scotus clearly do not refer to the same thing as the terms "state of affairs" or "circumstances" which I used in my earlier discussion of the cause for the extensive natural law's contingency. My "state of affairs" was a much broader term which was meant to encompass everything from basic features of the created order such as human nature to specific circumstances surrounding an act. Also, if Scotus's use of "*statu*" here were not more limited, these passages on universality would obviously conflict with his just-noted statements on dispensations in his discussion of marriage and bigamy. But on my interpretation of a more limited sense, there is no apparent conflict.

> principles, because they are seen by everyone to be very reasonable and good and observed justly by everyone, although they are not able to be demonstrated from them, like the former; and insofar as they are consonant with such principles, they are said to be of the law of nature, although not simply because they do not follow simply from them but from positive law. Others indeed are also consonant with such principles, but are so because the opposites of them are also consonant to them in that way, and so are able to remain standing with them, of which kind are ceremonies and customs, which are merely of the positive law only, like sacrifices and oblations and others of this kind.[32]

This reading would leave all three categories distinct: natural law strictly speaking would be strictly necessary; natural law in an extended sense would be consonant with natural law strictly speaking but would not follow simply (or, on my reading, would be contingent on circumstances for its content); and positive law would also be consonant with natural law but its opposites could also be consonant with it. Of course, this does not look like voluntarism.

But perhaps interpreters such as Allan B. Wolter[33] and Frederick Copleston,[34] who deny that Scotus actually represents voluntarism, are indeed

[32]*Rep.* IV, dist. 17, q. 1, n. 4: *"Alia autem sunt extra ista, quæ sic demonstrative concluduntur ex eis [i.e., principia practica nota ex terminis lumine naturali et omnia quæ immediate sequuntur ex eis], quæ multum sunt consona illis principiis practicis, quia ab omnibus sunt visa multum rationabilia, et bona, et ab omnibus juste servande, licet non possent ex eis demonstrari, sicut priora, et quatenus consona talibus principiis, dicuntur de lege naturæ, licet non simpliciter, quia non simpliciter sequuntur ex eis, sed ex Jure positivo. Alia vero sunt adhuc talibus principiis consona, sed ita quod etiam eorum opposita sunt eis ita consona, sicut illa, et ita stare possunt cum eis, cujusmodi sunt cæremonialia, et consuetudines, quæ mere sunt de Jure positivo tantum, ut sacrificia et oblationes, et hujusmodi."* Wadding v. 24, p. 275-276. See Hannes Möhle's reading of this passage: "Scotus's Theory of Natural Law," in *The Cambridge Companion to Duns Scotus*, ed. Thomas Williams (Cambridge: Cambridge University Press, 2003), 317.

[33]E.g., in *Duns Scotus on the Will and Morality*.

[34]E.g., in *A History of Philosophy*, vol. 2, *Mediaeval Philosophy: Augustine to Scotus*,

right. I believe the passage that provides the strongest support for these "mitigated" interpretations is probably the following, which occurs in a discussion of whether Adam's sin was the most gravest:

> Now, however, a sin which is only a sin because it is forbidden, is less of a sin formally, than that which is evil in itself, and not because it is forbidden. But now to eat of that tree, was not more a sin concerning the kind of act than to eat of another tree, but only because it was forbidden. But all sins which concern the ten commandments are formally evil not only because they are forbidden; but because they are evil, for that reason they are forbidden, since by the law of nature the opposite of what is pleasing was forbidden, and man is able to see through natural reason, which precept of these which are pleasing is to be kept.[35]

According to this passage, the sins forbidden by both the first and second tables of the Decalogue—i.e., natural law in both senses—are evil not only because they are forbidden but because they are evil in themselves. And "*quia mala, ideo prohibita*" seems like a straightforward denial of voluntarism.[36] Indeed, that line might be easily read as a direct response to Socrates's question to Euthyphro.

It seems to me that the majority of his texts lend support to a non-voluntarist reading of Scotus; that is, a reading in which the natural law in the extended sense is contingent because created reality is contingent

545-551.

[35]*Rep*. II, dist. 22, q. 1, n. 3: "*Nunc autem pecatum [sic], quod solum est peccatum, quia prohibitum, minus est peccatum formaliter, quam illud quod in se malum est, et non quia prohibitum. Nunc autem comedere de illo ligno, non plus fuit peccatum de genere actus, quam de alio ligno, sed solum quia prohibitum. Sed omnia peccata, quæ sunt circa decem præcepta, formaliter non tantum sunt mala, quia prohibita, sed quia mala, ideo prohibita, quia ex lege naturæ oppositum cujuslibet fuit malum, et per naturalem rationem potest homo videre, quod quodlibet præceptum ex illis est tenendum.*" Wadding v. 23, p. 104.

[36]Compare "*quia mala, ideo prohibita*" to Scotus's much-argued about statement that "everything other than God is good because it is willed by God, and not vice versa" (quoted in AW 16).

rather than being contingent directly on God's freely commanding.[37] This view, however, is not without its own problems. First of all, there is the explicit affirmation of voluntarism in the crucial discussion of natural law in *Ord*. III, suppl., dist. 37. Second, several of the key texts supporting a non-voluntarist reading are from the *Reportata Parisiensia*; although they represent Scotus's mature thought, they remain unexamined student reports, nor do we have a critical edition of them.[38] Because of these unresolved issues, I am not inclined to argue for a conclusive position regarding whether or not Scotus was a voluntarist.[39] Instead, I want to turn directly to the epistemological worry in order to see what bearing it has on this discussion.

Some Epistemological Worries Sketched

Suppose we accepted a strong voluntarism of the sort defended by Thomas Williams.[40] Even in this position there is natural law strictly speaking that is accessible to natural reason and that cannot be revoked by God. But according to Williams, virtually no moral applications follow from the first

[37]I realize that my discussion has proceeded as if there are only two basic positions: the voluntarist reading (which could be represented by Williams) and the non-voluntarist reading (which could be represented by Wolter and Copleston). Matters may not be quite that simple. One of the more interesting and compelling interpretations of Scotus is that of John Hare, e.g., in "Scotus on Morality and Nature," *Medieval Philosophy and Theology* 9 (2002): 15-38, which in some ways is certainly an intermediate position between Williams and Wolter. Ultimately, however, I worry that the "alternative routes" reading of natural law in an extended sense will be difficult to distinguish from merely positive law. Of course, I make no pretension to this comment being anything approaching an adequate response.

[38]Thomas Williams, "Introduction: The Life and Works of John Duns the Scot," in *The Cambridge Companion to Duns Scotus*, ed. Thomas Williams (Cambridge: Cambridge University Press, 2003), 10-11.

[39]A complaint I have about much of the secondary literature when conclusive positions are proposed is that a small number of texts are generally relied on by both sides. Scotus's corpus is much larger and it seems to me that it would be helpful, from a historical perspective at any rate, to give a more comprehensive reading rather than just playing the same several texts against each other repeatedly.

[40]E.g., "The Unmitigated Scotus" and "A Most Methodical Lover?: On Scotus's Arbitrary Creator." (A copy of the latter paper was given to me by John Hare.)

practical principle that one must not hate God. I take this to mean that with regard to most moral issues, God could have commanded us to φ or God could have not commanded us to φ. It does not seem to me that determining which is in fact the case will be accessible to natural reason alone since God's commanding is not to be explained by reference to the divine intellect, human nature, or anything else, according to Williams's definition of voluntarism. That leaves few options for how we might be able to know the moral commands. One possibility might be that God would have given us the relevant moral intuitions—or written the moral law on our hearts, to use traditional language. I am skeptical of this proposal since these moral intuitions would seem to be a relevant feature of that human nature in reference to which God's commanding is not to be based, according to Williams. That is, these moral intuitions would seem to be part of the circumstances upon which moral dicta may well be contingent even for a non-voluntarist.

What about divine revelation? Surely God could just simply tell us what moral commands he had in fact willed. Perhaps. But what reason would we have to think that God was telling us the truth? Presumably the idea that God would lie to us, especially about moral commands, seems repugnant. But I suspect that this is only because we already have a strong inclination to think that lying, at least in most cases, is morally repugnant. If we actually start with a largely unwritten moral slate, it seems we would have to grant that God could have decided that it was morally good to systematically deceive created beings on what the moral good was. Mark C. Murphy reports a similar argument made by Peter Geach specifically in reference to lying: "Now it is logically impossible that our knowledge that lying is bad should depend on revelation. For obviously a revelation from a deity whose 'goodness' did not include any objection to lying would be worthless. . . "[41] I think this argument can be extended to say that it is impossible for any knowledge of moral commands to depend on revelation from a deity whose goodness does not include an objection to lying. Therefore, to avoid falling into general moral skepticism, it seems to me that a principle proscribing deceit in at least most circumstances must be added to the principle that one

[41] Quoted in *An Essay on Divine Authority*, 34.

must not hate God as part of that moral realm accessible to natural reason apart from divine revelation.

This still does not seem enough, however. There is also the question of which god we ought to obey.[42] Obviously it is beyond the scope of this paper to substantially delve into apologetics; suffice it to say for now that I do not think either personal revelation or an attempt to figure out which god is the most powerful are sufficient means for identifying which claims of deity are true. Therefore, I think that one important way of deciding which god we ought to follow is by judging the gods' moral requirements. That we at least occasionally intuitively think that this is the right approach is illustrated by how compelling we find the story, reported by Robert Adams, about the men of the Purka clan who threw the whisk representing one of their deities into the nearest river, rather than comply with the deity's demand for human sacrifice.[43] Immanuel Kant makes a similar point when he discusses the command Abraham received to sacrifice his son Isaac. Kant concludes that the appropriate response would be to dismiss the command as being from a mere apparition rather than from God, as it is certain that killing one's son is wrong and God would not command one to do what is wrong.[44] In order to make this sort of judgment, however, there must be a fairly substantive body of moral knowledge that is accessible to us via natural reason, apart from divine revelation. Only once a judgment has been made about which deity we ought to follow can divine revelation be of help in filling the gaps in our moral knowledge.

It might be responded that the moral commands' fittingness or consonance with natural law strictly speaking would be sufficient for us to make this judgment, even if the moral commands were contingent on God's choosing between various alternatives that were all fitting in the appropriate sense. This could certainly go a long way. For example, any deity's require-

[42] Alasdair MacIntyre raises basically the same question using the same terms in "Which God Ought We to Obey and Why?" (*Faith and Philosophy* 3 (1986): 359-371), though he ends up working out his account differently than what I would have liked to see.

[43] *Finite and Infinite Goods: A Framework for Ethics* (Oxford: Oxford University Press, 1999), 282.

[44] For a discussion of Kant's response to Abraham's dilemma, see: Adams, 284.

ments that involved commands that were clearly not consonant with natural law strictly speaking could be rejected. In principle, however, it seems that two religious systems could offer two mutually incompatible ethical systems, both equally fitting with natural law strictly speaking. That is, if moral commands are contingent upon God's will insofar as God can choose between alternative morals as long as they are consonant with natural law strictly speaking, it must be the case that we could have two moral systems with at least one conflicting moral requirement. The difficulty arises if one deity demands one system and another deity demands another. It is difficult to see, in this case, how one could adjudicate between these two systems. Therefore, it seems that even this sort of consonance is insufficient. Natural law in an extended sense needs to be stronger.

Conclusion

The mitigationist reading of Scotus suggested earlier—namely, that natural law in its extended sense is contingent because created reality is contingent—largely escapes the epistemological worry. Some sort of proscription of deceit may need to be prior to even this sort of contingency, in order to escape worries about deceptive created orders à la skeptics' visions of demons and brains in vats and so on. Nonetheless, most of natural law could be contingent in this way without raising epistemological problems. But I fail to see how an unmitigated voluntarism like that of Williams can escape these worries.

The disagreement between the voluntarists and the mitigationists, however, may be less severe than it seems. The voluntarists want to emphasize the freedom and sovereignty of God, but even they concede that there is at least one practical principle from which even God cannot provide a dispensation. The mitigationists want to emphasize the rational accessibility of moral law, but even they are willing to concede that there are at least some commands that are dependent solely on God's willing. At the heart of the disagreement is not really the question of what are the appropriate categories, but rather how best to identify the proper domain of each category.

Works Cited

Adams, Robert Merrihew. *Finite and Infinite Goods*. Oxford: Oxford University Press, 1999.

Copleston, Frederick. *The History of Philosophy*, vol. 2, *Mediaeval Philosophy: Augustine to Scotus*. London: Burns Oates & Washbourne, 1950.

Duns Scotus, John. *Opera omnia*. ed. Luke Wadding. 26 vols. Paris: Ludovicum Vivès, 1891-1895.

Hare, John. "Scotus on Morality and Nature." *Medieval Philosophy and Theology* 9 (2002): 15-38.

Kant, Immanuel. "Grounding of the Metaphysics of Morals." *Practical Philosophy*. ed. Mary Gregor. Cambridge: Cambridge University Press, 1996.

Kovach, Francis J. "Divine and Human Beauty in Duns Scotus' Philosophy and Theology." *Studia Scolastico Scotistica* 5 (1972): 445-459.

MacIntyre, Alasdair. "Which God Ought We to Obey and Why?" *Faith and Philosophy* 3 (1986): 359-371.

Möhle, Hannes. "Scotus's Theory of Natural Law." *The Cambridge Companion to Duns Scotus*, ed. Thomas Williams, 312-331. Cambridge: Cambridge University Press, 2003.

Murphy, Mark C. *An Essay on Divine Authority*. Ithaca: Cornell University Press, 2002.

Plato. “Euthyphro.” *Plato Complete Works*. ed. John Cooper. Indianapolis: Hackett, 1997

Stratenwerth, Günter. *Die Naturrechtslehre des Johannes Duns Scotus*. Göttingen: Vandenhoek & Ruprecht, 1951.

Williams, Thomas. “A Most Methodical Lover: On Scotus’s Arbitrary Creator.” *Journal of the History of Philosophy* 38 (2000): 169-202.

———. “The Unmitigated Scotus.” *Archiv für Geschichte der Philosophie* 80 (1998): 162-181.

———. “Introduction: The Life and Works of John Duns the Scot.” *The Cambridge Companion to Duns Scotus*, ed. Thomas Williams. Cambridge: Cambridge University Press, 2003.

Wolter, Allan B. *Duns Scotus on Will and Morality*. Washington DC: Catholic University of America Press, 1997.

Are There Aristotelian Substances?

D.R. Foster
Oxford University

There is a broadly Aristotelian conception of substance, supported in various forms by Aristotle himself, and later by Ayers and Wiggins, which takes substance as the ontologically basic, absolute, persistent, unitary, irreducible, and individual subject of all predicates. This paper interrogates this account in light of the everyday intuitions which underlie it—metaphysical intuitions about the persistence of particulars over time, as well as semantic and linguistic intuitions regarding the practices of individuation and predication. After surveying some of the relevant literature on Aristotelian substance-theory, Foster argues that in its current state, the Aristotelian conception of substance is unattractively dualistic, as well as metaphysically extravagant. As an alternative, the paper argues for a modified "bundle-theory" of substance, which holds that Aristotelian substances can and should be understood as nothing more than structured matter with manifest "bundles" of properties. On this basis, Foster attempts to demonstrate that our semantic and logico-linguistic intuitions can be fruitfully explained and vindicated by a bundle-theory, while remaining agnostic about the richer metaphysics suggested by the Aristotelian theory.

The question I set out to answer in the title of this paper is a question about whether *one kind* of account of substances (the Aristotelian account) is the *right* account. But in attempting to answer that question, we will need to engage—directly or indirectly—with even more fundamental questions, such as whether there is such a thing as substance, and whether we need a notion of substance to understand reality. But if we are to suggest answers to these questions, we will of course need to fix the definition of what we understand by 'substance' in the first place. Historically, this has proven to be extraordinarily difficult; the (purported) very nature of substance is such that reductive definitions of it are hard to come by. As a result, many

of the orthodox philosophical treatments of substance give accounts of it which are: stipulative and taxonomic (as some critics would argue of the early Aristotelian conception); wholly or reservedly skeptical (with Hume's position typifying the former and some of Locke's comments the latter); or else transcendental or semi-mystical (as, arguably, Aristotle's is in the *Metaphysics*). The difficulty of arriving at a more essential definition of substance is surely bound up with its semantic and linguistic inextricability from similar "family circle" terms (such as the terms 'fixed,' 'definition,' 'nature,' and 'essential,' and even the terms 'it' and 'is,' all of which are used in the foregoing paragraph). The fact of this semantic tangle is understood and acknowledged even by positive substance-theorists (such as Wiggins in his preamble). Still it may lead others to conclude, on Quinean grounds, that the notion of substance is hopelessly confused, in that it cannot be given an adequate explanation without resorting to concepts and terms with which it is mutually, semantically, and logically interdependent. The Quinean connection is perhaps a new gloss for an old conclusion of the kind held by the classical empiricists: namely, that the concept of substance is at best a useful fiction, a mind-dependent discrimination unsusceptible to empirical confirmation or epistemic justification. I will not go quite this far in this paper; but I *will* argue that the Aristotelian conception of substance, and conceptions like it, are superfluous. That is, I hope to suggest reasons to believe, via a kind of bundle theory of substance, that we can retain all the important epistemic, semantic and logico-linguistic features of Aristotelian substances while remaining agnostic on the deeper metaphysical questions.

In the *Categories*, Aristotle takes substance to be chief among the categories of being, along with quantity, qualification, relativity, spatio-temporal location, orientation, and possession and action. Aristotle gives "rough" examples of these categories:

> examples of substance are man, horse; of quantity: four-foot, five-foot; of qualification: white, grammatical; of a relative: double, half, larger; of where: in the Lyceum, in the marketplace; of when: yesterday, last-year; of being-in-a-position: is-lying, is-sitting; of having: has-shoes-on, has-armour-on; of

> doing: cutting, burning; of being-affected: being cut, being-burned. (Aristotle 1963: 1b25)

Substance is uniquely the subject of predication for all the other categories, and falls into two types. *Primary substance* is "that which is neither said of a subject nor in a subject, e.g. the individual man or the individual horse." In addition, there are *secondary substances* which are the classes, kinds, or to use Aristotle's term, the genera under which the primary substances fall (1963: 2a11). As substances are the ultimate subject of predication, "all other things are either said of the primary substances as subjects or in them as subjects. So if the primary substances did not exist it would be impossible for any of the other things to exist" (1963: 2a34). Another way of thinking about this is to say, with Ayers, that:

> [In Aristotelian ontology] only an individual horse can be the subject of the substantival predicate 'horse,' whereas the subject of the predicate 'brown' is not an individual brown but, for example, a horse. If such an individual as *this brown* can indeed be identified and made the subject of predication... it exists parasitically on a substance, e.g. for just as long as this horse is brown... Non-substantial things are the modes of the being of substantial things. (Ayers 1991a: 18)

In addition to their ontological priority, substances are: (i) "individual and numerically one," not validly separable into more basic units (Aristotle 1963: 3a29); (ii) absolute, not admitting of degrees of substance-hood (Aristotle 1963: 3b33); and (iii) without contrary but nevertheless able to *receive* contraries while persisting. This is to say that one and the same substance—an individual man—"becomes pale at one time and dark at another, and hot and cold, and bad and good." (Aristotle 1963: 4a10)

It is also important here to note exactly what substance *isn't*. It isn't, for one thing, congruent with our ordinary conceptions of 'object,' 'thing,' or 'matter.' Elsewhere, Aristotle flirts with a corpuscularian or atomistic conception of concrete particular things as the conjunction of matter and form,

but in the *Categories*, substances are irreducible to structured matter. Aristotle found the theory "that the elements of the universe are immutable material atoms... contrary... to the commonsense of mankind, which assumes that many sensible bodies, as such, are substances with irreducibly unitary natures" (Ayers 1991a: 24). Thus, the Aristotelian ontologist needn't start from *physical* individuals such as immutable material atoms, or even *conceptual* individuals such as simple ideas and sense impressions, to arrive at that which is nevertheless a numerically one and individual *substance*, such as a person or animal or plant. We can note here that insofar as this doctrine at least *seems* to entail a kind of dualism about concrete particulars like persons and animals, a dualism which posits a difference between particular *qua* structured matter and particular *qua* substance, then it is a doctrine at odds with the prevailing trends in analytical philosophy, which are reductively-physicalist. And, while substance-matter dualism in the case of certain particulars doesn't seem to *directly* entail other, more dubious dualisms, it makes for an odd sort of theoretical relic amid a philosophical order which prefers that descriptions of things be empirical wherever plausible.

Ayers and Wiggins each offer nuanced but broadly Aristotelian theories of substance, springing from similar philosophical concerns about vindicating our common sense intuitions about persistent particulars, and with explaining and justifying everyday practices such as individuation and predication. But, were it the case that these concerns could be allayed by an ontology which did not rely on Aristotelian substances (that is, were it the case that our linguistic, semantic, and epistemic practices could be vindicated without appeal to substances), then, it seems, the Aristotelian position would look far less attractive. We can now briefly investigate this possibility.

Ayers sets himself to defend six principles of substance, as opposed to "[a]ttributes, such as a thing's redness or squareness, and events, actions, or processes, such as walks, walking, thunderstorms, and the like" which are themselves not substances but modes of substance (Ayers 1991b: 69):

(1) Substances are the ultimate subjects of predication, and therefore the

only beings with independent existence.

(2) Substances are real unities (both natural and logical).

(3) Substances are material. Individual substances are distinguished from one another at any one time by their matter.

(4) Substances exist all at once, and exist through time, or endure. (Events, in contrast, take time or unfold).

(5) Substances are active, the ultimate sources of change. Their underlying natures or essences are the ultimate principles of explanation.

(6) Only substances fall into true natural kinds, and every truly individual substance is a member of a natural kind. (1991b: 70)

Call these six principles *A*. Now, let us presuppose an alternative to the Aristotelian view, a view that holds that substances are merely individual bundles of spatio-temporally coordinate properties manifest by some structured matter. Call this theory *B*. This view typically rejects the intuition that there is a metaphysical substratum above and around which modes cohere; but the version of bundle theory I argue for merely remains agnostic about the metaphysical question and seeks to explain using a more austere ontology our practices of individuation, predication, etc. This version claims that we need no more than bundles of spatio-temporally coordinate properties manifest in an extended substance to explain our intuitions about substance. So how does such a view deal with the propositions of *A*?

One can notice from the outset that only (3) appears to be wholly consistent with a non-Aristotelian account of substance. That is, the discernability of individuals would hold even if instead of individual substrate substances one presupposes bundle substances, since part of the stipulated theoretical definition of substance in *B* is that substance be material manifest.

The other principles of *A* we can deal with in turn. (2) and (6) are aspects of the familiar Aristotelian claim about the wholeness and exclusivity of substance. According to this claim, *primary* substances are natural and logical wholes and constitute as a class *secondary* substances or natural

kinds. The claim issues from the observation that substantive individuals, as opposed to what Ayers takes to be constructed individuals (individuals like the ones presupposed by *B*), are naturally *given*, such that

> the term 'horse' simply *classifies*, or gives a general name to, natural individuals. . . [Whereas] the adjective 'brown' . . . *characterizes*, not browns themselves, but brown objects. Particular browns are sliced out of reality by the concept *brown* or, if you like, by the nominalized predicate, 'brown' or 'brownness.' Such entities are *abstract* particulars, mere aspects of things. Their individuality is posterior to their individuation by us, or by our concepts. (Ayers 1991b: 71)

The "boundary conditions" of many non-substances are determined pragmatically; that is, they are determined by our concepts and their usefulness in making distinctions. "It is our concept that determines that the smoke is an effect of the fire, whereas the flame is a part of it. . . By contrast, a horse is by its nature a whole, its undetached hoof is by nature a part of that whole" (1991b: 74-75).

But even if this is the case, then *logically* speaking there is no necessity about an individual's substance-hood—that is, about the logical and natural unity of substances claimed by (2). It might have been the case that horses were not substances but machines run by teams of smaller animals. Similarly, it might sometimes be the case that a particular constructed of a spatio-temporally coordinate bundle of properties (say a bundle that describes a fictional character) will actually name a substance. As both these facts are contingent, it seems substance-hood has the same general modal status as non-substance-hood. What this suggests is that even if substances are natural wholes in some coherent and observable way, it does not follow that they are necessarily logical wholes. It is not nonsensical to talk about a purported substance turning out to be a constructed bundle, or vice versa. In practice this means that whatever inductive grounds we may have for believing that the contiguous perceived unity of a bundle of properties constitutes a substance in the Aristotelian sense, our belief will be just that:

inductive. There is nothing in the contents of our perceptions that validates the move from natural or sensible unity to logical unity, and so we can always talk about substances and non-substances in the same language with the same modality. Similarly, if primary substances aren't *necessarily* primary substances—if we cannot validly move from their apparent empirical unity to a logical unity—then of course we cannot hold that, necessarily, only true substances can constitute secondary substances or natural kinds. Simply, there is nothing to stop one from coherently recognizing a kind that subsumes every individual with a particular bundle of properties; and since it is a contingent fact whether that bundle of properties names a substance, we can't be logically sure that that kind isn't a natural kind. This, it should be fairly obvious, is merely a rather technical gloss of Humean epistemic skepticism about individuation.

Ayers maintains that particular modes or property bundles

> are abstractions sliced out of reality by general concepts. It is therefore unsurprising that modification of the predicate is liable to breed fresh individuals. Particular substances, however, are naturally discrete individuals whose given individuality is prior to their being identified and classified by us. It is therefore unsurprising that mere addition to the classificatory predicate or general name we employ will *fail* to breed fresh individuals. (Ayers 1991b: 82)

But of course this is a metaphysical claim; there is no *speaker-relative* guarantee that substances will be picked out as substances, non-substances as non-substances. We need look no further than Kripke-style puzzles about belief to see this. Even if names rigidly designate in a metaphysical sense, as Kripke would have it, this does not imply that we can have good knowledge of which names are 'logically pure' and designate natural kinds, and which names are merely abbreviated descriptions which take bundles of properties as their referents. What I mean is that there is no obvious and equally rigid *epistemic* constraint on holding beliefs about individuals. I may hold both the belief that Clark Kent cannot fly and that Superman can fly, that Kent is

awkward and quiet and that Superman is charming and heroic. I can hold these pairs of beliefs *consistently* because *logically* I have constructed individuals out of bundles of properties that I have presumably observed as materially extended. And *empirically* all I have are discrete sense impressions à la Hume. Even if we go so far as to admit over and above the skeptic that these sense impressions *are* of contiguous and *objectively existing* materially extended bodies from which I have constructed logical individuals, there is no empirical promise that I can come to know that the body to which the Kent bundle belongs and the body to which the Superman bundle belongs are one and the same body.

Similarly, we can go beyond Kripke-type cases and imagine an instance in which a property bundle and accompanying name seem to pick out an individual underlying substance, but actually pick out multiple substances. As opposed to the fantastical example used earlier of a horse-like machine run by many smaller animals, I have a more realistic example in mind: the Portuguese Man-of-War. It might realistically be the case that for all our casual empirical investigation, our knowledge of a particular Man-of-War would give us grounds to take it as a substance and subsume it under a sufficiently broad genera or natural kind with other jellyfish-like individuals. But, as the Australian Museum's website notes, the Portuguese Man-of-War "is not a single animal but a colony of four kinds of highly modified individuals [polyps]. The polyps are dependent on one another for survival." Now, the four kinds are spatio-temporally contiguous, seem to have the same existence conditions, and are liable to coherent predication. But science has good reason to believe they are nevertheless *four kinds*. Nevertheless, in our everyday usage, the name for the natural kind is singular rather than plural, and we use a singular ostensive pronoun "that" to pick out a particular member of the natural kind. But are we picking out an individual substance? In this instance it is certainly empirically dubious that we are picking out an individual unity, but it is also unclear exactly what the *metaphysical* fact of the matter would be like even if there *were* a metaphysical fact of the matter.

But there is perhaps an even more confusing case, where the epistemic and metaphysical facts of the matter are thoroughly obscure. Suppose we

have good experiential grounds for believing that the name "Dr. Jekyll" and a sufficiently exhaustive associated property bundle pick out the same material body, e.g. Dr. Jekyll, and thus conclude that underneath the body and the bundle is a Jekyll substance. Now, whether or not we actually go out and get knowledge of it, we can suppose that it will be a potentially knowable fact about the world that there is a man, Mr. Hyde, whose body is identical with Jekyll's. If we suppose that the Jekyll body and its property bundle cohere in a single substance, does the Hyde property bundle—radically different from the Jekyll bundle but ascribable to the same body—cohere in the same substance? If, as my professor Peter Caws suggests, we should think of Jekyll and Hyde cases not as cases of 'multiple personality' but of 'multiple *subjectivity*,' then what are our *metaphysical*, let alone our *epistemic* options in characterizing the substances and properties involved? Can we maintain an identity between Jekyll and Hyde along Leibnizian grounds by positing that Jekyll has all its own properties, in addition to the properties of being Hyde, and that Hyde has all its own properties in addition to the properties of being Jekyll? It seems this construction already entails that there are numerically two individuals. We might then explicitly say that it is one and the same individual who has the properties of Jekyll and Hyde. But besides the fact that Jekyll and Hyde themselves will be unlikely to assent to this utterance, what then do 'Jekyll' and 'Hyde' refer to if not individuals?

The above discussion aims to point out where my bundle theory *B* and Ayers's Aristotelian theory *A* differ with respect to the question of the unity of substances. I hope to have shown that the logical and epistemic facts about certain individuals do not compel us to adopt *A*, but rather the opposite. My arguments have overlapped somewhat in dealing with claims made in (1), (4), and (6) of *A*, claims about the independence, endurance, and change-bearing nature of substance. I will now deal directly with those claims, and aim to show how they lose their force with respect to a positive bundle-theoretic account of what it is we do when we individuate things as substances.

A asks us to accept that substances, properly conceived, are the ultimate subjects of predication and the ultimate source of change, but that they

themselves exist all at once as whole and persist over time. Substances are what Wiggins calls 'continuants.' What is *B*'s counterproposal? *B* rejects the notion that there is any substratum or bare subject stripped of its accidents which we can call the ultimate subject of predicates. This rejection stems foremost from the familiar difficulties with demarcating the boundary between essential and accidental properties. The classical anti-Aristotelians have enumerated several such difficulties. We might add to this enumeration by taking the modal argument above as roughly suggesting that if essentialism requires that essential properties are held necessarily, it is difficult to sustain the argument that substances are *essentially* substances, since it is difficult to argue that that they are *necessarily* substances.

Instead of supposing a substratum, *B* argues along the lines that substances are something like databases, or more or less systematized and interrelated bodies of data. When we predicate on a substance, we 'call up' a particular datum and recognize it with respect to the other data in the database. Supposing Socrates to be a paradigm substance, when we predicate "has a beard" on Socrates, what we are actually doing is looking at one datum in the context of the Socrates database. The database, or the bound totality of known facts which we have reason to believe cohere in one spatio-temporally contiguous body, takes on subjectivity with respect to any particular fact. 'Socrates' *names this database*, which is contained in our brains; naming the database from which a particular datum comes gives us context for understanding the significance of that datum. That *Socrates* has a beard is significant in a way distinct from the way it is significant that *Plato* has a beard, because the datum 'has a beard' is predicated with respect to subject-sets that contain different data.

Such an account worries Wiggins because it fails to accommodate the endurance of substance and the absoluteness of certain of its features. He says that

> Unlike sets or classes [(we can add databases for our purposes)], continuants can gain and lose elements. The terms of the identity questions that we are concerned with. . . stand for these continuants themselves, John Doe, say, not for phase-thick lamina-

> tions of their four-dimensional counterparts, John Doe-when-a-boy, or for infinitesimally thin time-slices of these held fast between perceptual impressions, John Doe-as-caught-sight-of-there-then. And, by the same token, the phased sortals like 'boy' or 'old man' that we encounter in English never denote 'phases' of entities or (if that were different) the entities themselves frozen at an instant. They denote the changeable changing continuants themselves, the things themselves that are in these phases. (Wiggins 32)

But of course, it is only the case that sets and classes cannot gain and lose elements because that property is stipulated in their formal definition. We might just as well define the notion of a 'continuant' or a 'substance,' then, as a set or class that *can* gain and lose elements. What we would name, then, would be the database itself. *B* has it that when we name substances we actually name databases, or to put it another way, we bracket off, with our language and our cognitive capacities, substances. We bracket 'vertically' by type or kind, and 'horizontally' by individual or particular. But in allowing for substances to take on change, we are actually naming not the bunch of properties we determine to include in the brackets, *but the brackets themselves*.

This means that, according to *B*, even if there are substances, these substances have accidental and maybe essential properties. And the fact remains that I can, if I choose, bracket off certain of these properties, give my bracketing a name, and ascribe further predicates to the bracketed properties. So when I bracket off only those perceivable properties that constitute a statue of a Biblical figure, and name that bracketing 'Goliath,' I can ascribe to Goliath (the bundle of properties) the further predicate that Goliath is made out of Lumpl.[1] That I can do this is surely due to facts about my language, my logic, and my cognitive capacity for abstract reasoning. But that I *actually* do this is surely due to pragmatic facts, facts about contexts in which it is a useful linguistic practice to construct a logical object by nam-

[1] See Gibbard (1975)

ing a bracket of sensible properties. There may a metaphysical fact of the matter about my linguistic practice, but it does not follow that my linguistic practice is therefore false or mistaken.

In short, it appears that, even with a healthy skepticism about the epistemology of substance, and an austere ontological explanation of what substance is, we are still in a position to offer justification for the commonsense intuitions that spawned our worries over substance in the first place. If my arguments have been successful, this result will suffice to show that the Aristotelian conception of substance is at best analytically cumbersome and superfluous, and at worst metaphysically wrong.

Works Cited

Aristotle. *Categories and De Interpretatione*. trans. J.L. Ackrill. Oxford: Clarendon Press, 1963.

———. Metaphysics. Translated by Hugh Lawson-Tancred. London: Penguin, 1998.

Ayers, Michael. *Locke*. London: Routledge, 1991.

———. "Substance: Prolegomena to a Realist Theory of Identity." *The Journal of Philosophy*. 88.2 (1991a): 69-90.

"Bluebottle or Portuguese Man-of-War." *Australian Museum*. 2003. <http://www.amonline.net.au/factsheets/bluebottle.htm>.

Gibbard, Allan. "Contingent Identity." *Journal of Philosophical Logic* 4 (1975): 187-221.

Hume, David. *Enquiry Concerning Human Understanding*. ed. Tom L. Beauchamp. Oxford: Oxford University Press, 1999.

Kripke, Saul. "A Puzzle About Belief." *Meaning and Use*. ed. A. Margalit. Dordrecht: Reidel, 1979.

Quine, Willard Von Orman. "Two Dogmas of Empiricism." *From a Logical Point of View*. 2nd ed. Cambridge: Harvard University Press, 1961.

Salmon, Nathan. "How to Become a Millian Heir." *Noûs*. 23.2 (1989): 211-220.

Wiggins, David. *Sameness and Substance Renewed*. Cambridge: Cambridge University Press, 2001.

What is a Mood?

Al Prescott-Couch
Columbia University

The idea of "moods" is ubiquitous in our everyday lives; theories about mood attempt to capture these commonplace intuitions, while providing a clear and philosophically satisfying definition of the phenomena. In this paper, Prescott-Couch uses a set of criteria suggested by Eric Lormand to evaluate competing theories of mood. Focusing on three of the most plausible accounts—Generalization, Functional Component, and Higher Order Functional State theories—the paper argues that the third view, held by P. E. Griffiths, is the most plausible. Prescott-Couch then entertains objections to the functional approach, paying particular attention to the phemenological worry—that is, that this type of theory fails to grasp at our subjective experience of moods.

We use the word "mood" in a variety of ways. I can be in a good mood after taking an exam, and I can be in the mood to start my philosophy paper. A sentence can be in the indicative mood, a country can be in an isolationist mood, and if I'm particularly anti-social or irritable, others might describe me as moody. This paper will be concerned with the type of "moods" defined as nonspecific and indeterminate affective states. Depression, anxiety, ennui, irritability, elation, and levity are all examples of moods.

Eric Lormand proposes a nice way to distinguish between the moods we wish to discuss here and other uses of the term mood. The moods we're talking about can be described by locutions of the form "A is in a __ mood" but *not* by locutions of the form "A is in the mood to __" (Lormand 385-6). Thus, a state like a "beer-drinking mood" cannot be a mood of the sort we're looking for because it could be redescribed as "the mood to drink beer." But things like "an anxious mood" or "a good mood" *are* the type of phenomena

we're looking for, since neither of these can be redescribed as being in a mood to do anything. This "criterion" is just a way of illustrating that the moods we wish to discuss are abiding psychological conditions rather than simple desires (being in the mood to drink beer is basically equivalent to desiring to drink beer[1]). It also points to another feature of moods that I will discuss shortly: their lack of particular intentional objects.

In everyday talk, moods are commonly conflated with other affective states such as emotions and character traits. This is because many mood words are also used to describe emotions or character traits. For example, one can be anxious (mood), anxious about something (emotion), or an anxious person (character trait). However, we should not let common language deceive us into thinking that moods and emotions are the same, or even similar, phenomena. There are fundamental differences between moods and emotions, although there is some disagreement over exactly what those differences are. In "Moods, Emotions, Traits," Paul Ekman details a few of the many criteria psychologists use to differentiate the two: duration (moods typically last longer), facial expressions (many emotions have unique facial expressions, while moods do not), and causal antecedent (people are usually aware of what caused an emotion, but it is frequently unclear what caused a particular mood). Some psychologists simply view moods as dispositions to particular emotions (for example, irritability is a disposition for anger, depression a disposition for sadness, etc.) (Ekman 56-7). Other psychologists take a more dramatic approach. Richard Lazurus claims moods "are products of the existential background of our lives. This background has to do with who we are, now and in the long run, and how we are doing in life overall" (Lazurus 84).

Some of the preceding criteria for distinguishing between emotion and mood seem to be trivial (who cares if one set of states has corresponding facial expressions and another doesn't?), while others seem unproven (one

[1]It may be necessary to distinguish two types of desires, which we can call "volitive" and "appetitive." Volitive desires are goals. For example, I (volitionally) desire to study. Appetitive desires are more like cravings. They are not grounded in reasons. An appetitive desire is something akin to "being in the mood to do something."

should not *assume* that a mood is merely a disposition to feel certain emotions). Philosophers have had more success. Eric Lormand's "Towards a Theory of Mood" presents three phenomenologically compelling criteria for establishing a relevant distinction between mood and emotion: intentionality, pervasiveness, and causal explanatory structure. These three properties of mood must be explained by any theory of mood in order for it to be a compelling theory. I will briefly describe these three criteria, and then provide arguments why they are compelling.

First, moods lack obvious intentional objects. They are not *about* anything. When you are irritable, you are simply irritable. You are not irritable *at* someone. When I am depressed, I am not depressed about anything in particular. Emotions, however, *are* directed towards states of affairs and objects under certain representations. To be angry is to be angry *at* someone or *about* something. Lormand proposes that any theory of mood must explain why moods at least appear to lack intentional objects while emotions do not. This, he calls the Intentionality Condition.

Second, moods have profound and far-reaching effects on a wide variety of our mental states. Emotions usually do not. A change in mood from depression to euphoria has much wider psychological ramifications than a change from being angry at my roommate to liking him. Moods tend to color everything one sees in the world, explaining the abundance of visual metaphors associated with moods. Moods are described as "filters." When you're in a good mood you "see the world through rose-colored glasses," and when you're depressed, "everything looks dark." Unlike emotions, moods cast *every* object of experience under a certain light. Any adequate theory of mood must explain why moods have such pervasive effects. Lormand calls this the Pervasiveness Condition.

Third, moods explain actions in a very different fashion than do goals, desires, and emotions. In our normal explanations of action, we reference the motivations of the agent in conjunction with his beliefs and show how these rationally combine to form an intention to act. For example, if we are trying to explain why I yelled at my sister, we might reference my belief that she had thrown out some of my books and my belief that if I yell at her

she will not throw out my books again, in conjunction with my desire for her to not throw out any more of my books. This combination of beliefs and desires makes the action intelligible. However, when we explain actions by reference to moods, reasons and beliefs do not enter the account. If I state that "I yelled at my sister because I was in a bad mood," I am not citing reasons and desires that conjoin rationally to explain my intention. Mood explanations are merely *causal* explanations. Explaining an action by citing a mood does not provide any motivation for the action.

Lormand argues that, similarly, the explanations of moods themselves do not reference practical or goal-directed reasons. Moods are not grounded in goals and beliefs as emotions are. While it may be rational to feel certain emotions, it is not "rational" to feel any moods. For example, if it is rational to dislike someone who commits murder, and Joe commits murder, then it may be rational to be angry at Joe. However, it would not be rational to be irritable or anxious in general. It does seem to be true that emotions somehow *cause* moods (being angry at Joe can put you in a generally irritable or anxious mood), but the transition is not rational, as it does not employ reference to beliefs or desires; it is merely causal. This also explains why moods can be caused by purely physiological states like headache or stomach pain. It would be bizarre to say that a headache caused me to be angry at my brother. But it is commonplace to say that a headache made me irritable (put me in a bad mood). Explanations of causes *of* moods and explanations of actions caused *by* moods are both causal (they do not appeal to beliefs and desires of the agent) while those of emotions are not. This is the Explanation Condition (Lormand 388-91).

These three conditions stipulate the properties of mood that any theory of mood has to account for. The rest of this paper will examine the strengths and weaknesses of three different theories of mood. While there are a number of other significant theories of mood—particularly theories of Duration and Disposition—I will not examine them here in order to spend more time with the three theories that I believe are most plausible; namely, generalization, mood as a functional component, and mood as higher order functional state. I will argue that among the current theories, the thesis proposed by

P.E. Griffiths in "Folk, Functional, and Neurochemical Aspects of Mood" best captures of the nature of mood and that the functionalist program has the best chance of future success in the area. I will also raise some questions regarding whether or not the inability of functionalism to capture the qualitative aspects of our experience of moods is a serious liability.

Generalization Theory

A prominent theory of mood holds that moods are generalized emotions. Here, moods differ from emotions only in degree, rather than in kind. According to this view, the difference between moods and emotions lies only in the difference of their intentional objects. While emotions are *about* specific people, objects, or states of affairs, moods are about "general" objects such as "the world as a whole," "life itself," "my existence," etc. Robert Solomon, among others, has held this view. This theory of mood is convenient in that it enables moods to be analyzed in the same fashion as emotions (according to belief and desire). It also explains why the effects of moods seem so pervasive. Moods color our experiences of all objects because the individual objects are recognized as part of the world as a whole (the general object), and so the valuations of the world as a whole fall onto the individual objects that make it up.

However, this account is unsatisfactory for a number of reasons. First, it does not take into consideration that, unlike with emotions, we do not *reason* ourselves into or out of moods. It fails the Explanation Condition. Second, there is a question as to the status of objects like "the world as a whole." Can they be actual intentional objects? Even assuming that we can have beliefs about "the world as a whole," how similar are these beliefs to paradigm beliefs with typical intentional content? Moreover, our moods change with a fair amount of frequency. Do all our beliefs about the world change this quickly and often? Similarly, how do we explain these quick changes in a substantial number of our beliefs? All these questions should cast doubt on the Generalization Theory.

Third, as Eric Lormand points out in "Towards a Theory of Mood," if

moods really are just beliefs and desires about "the world," then moods should change when these beliefs and desires change. And this is just not the case. If someone is truly depressed, she will not be cured by being shown all the wonderful things in life. Many depressed people can assent to the proposition that the world really is a good place, but this does cure them of their depression. Likewise, as Lormand points out, "I may believe, sincerely and strongly, that the world is unspeakably cruel, and I may desire, sincerely and strongly, that it not be that way, but I am not in a depressed mood if I don't let these states (or something else) *get me down*" (Lormand 393).

Finally, the most serious flaw in the account is that it may be circular. In "A Computational Theory of Mood," Laura Sizer argues that "describing moods as having general objects like 'everything,' 'things in general' or 'the world as a whole' does not yield explanations or means of identifying the states distinct from the stipulation that the subjects are in those states" (Sizer 751). To illustrate her point, Sizer proposes the following scenario: Suppose you have two people, Marvin and Esther, each in a different mood state. Marvin is melancholic (M1) and Esther is ebullient (M2). How does the generalization theory explain the difference between their two moods? Marvin is melancholic (about everything) and Esther is ebullient (about everything). We see that Marvin is melancholic because he is melancholic and Esther is ebullient because she is ebullient. In other words, "the only thing doing the explanatory work is the stipulation that the subject is in a particular mood" (Sizer 750).

Mood As a Functional Component: Eric Lormand

Eric Lormand begins his analysis of mood in "Towards a Theory of Moods" by noting that with the failure of the Generalization Theory, the only way to account for mood's lack of intentional content is that moods *are* genuinely non-intentional. Thus, they cannot be analyzed into beliefs and desires. He argues that this also explains why moods do not share the explanatory structure of emotions. Since moods do not involve beliefs or desires, they

cannot enter into rational calculation. Thus moods can only have causal explanatory properties. They are immune from practical justification (one cannot give "reasons" for being in a bad mood in the same way that one can for being angry), and they cannot be used to explain actions as rational (they can only explains actions causally) (Lormand 394).

Lormand then makes the distinction between *active* beliefs and desires and *latent* beliefs and desires. Active beliefs and desires are those that are immediately necessary for the explanation of an agent's action. (Here, "action" is construed broadly to include things such as where to focus attention and one's theoretical or practical reasoning, in addition to overt physical behavior.) Active beliefs and desires need not be "conscious" of course. Latent beliefs and desires are those possessed by an agent but not immediately necessary for an explanation of her current actions. These are the agent's "stored" beliefs. (Lormand uses the term very loosely because he does not want to commit to a realist position about beliefs and desires.) Lormand admits that exactly how beliefs and desires transform from being active to latent is a complete mystery. He knowingly posits a black box with a homunculus to explain the transitions.

Lormand hypothesizes that there is a functional component of the mind *M* that is functionally positioned between the "control unit" (the part of the mind guiding behavior, directing attention, reasoning into new intentional states, etc.) and the store of latent intentional states. This functional component hampers or facilitates different certain classes of beliefs and desires. Different states of *M* correspond to our different moods. To summarize, according to Lormand, a mood is a state of the functional component of the mind that inhibits or facilitates access to certain latent intentional states.

One of the counter-intuitive consequences of this theory is that moods do not directly affect the contents of our beliefs: "Moods are not so much in the business of *changing* one's beliefs and desires as in... helping to determine one what beliefs and desires one *dwells*" (Lormand 398). Since this seems like a problem (surely moods affect our beliefs about people, objects, etc.!), Lormand provides an account of how moods can change beliefs and desires indirectly. According to his view, since moods alter which features

of a particular object are attended to and which beliefs and desires function in practical reasoning, one's assessment of a person, object, or state of affairs will be affected by one's mood (since the mood influences the information one uses to make the assessment). For example, if a friend of mine seems more annoying when I'm in a bad mood, it's because I become more aware of the annoying aspects of her personality than I otherwise would be. None of my beliefs about her individual characteristics have changed, but my assessment of her as a whole has changed because the information I'm using to make the assessment changes. Likewise, to explain why good moods lead us to the belief that "the world as a whole is wonderful," we can say that the mood influences which evidence we use to draw our conclusions; since a good mood leads to a focus on positive aspects of the world, we induce that the world as a whole must be super.

While this theory is certainly stronger than the previous theories, I believe Lormand's account of how mood affects belief flies in the face of our phenomenological evidence. Let's take the example of my friend again (the one who is quite funny, but who annoys me when I'm irritable). When I am in a bad mood, it is not the case that I simply focus on things about my friend about which I already have negative feelings. It's that I construe each of her individual behaviors differently. Jokes I would have found funny, I now find grating. Behavioral quirks which I usually find cute, now appear irksome. Lormand's account simply does not do justice to the phenomenology of mood.

Lormand's theory has a few other consequences that we should discuss, because they will clarify other questions regarding mood. Since *M* is a functional component, one is always in an *M*-state, and hence always in a mood. Even when we're going about our everyday activity with our average disposition, we're still "in a mood." This is not how we commonly use the term "mood." "Being in a mood" usually implies being in a state different from our average disposition. A mood is frequently said to "distort" or "color" our vision of the world. However, common language is probably not a particularly helpful guide in these matters, and the dispute seems to be merely verbal. Is our average way of mediating between latent and active

beliefs and desires considered a "mood"? If yes, then we're always in a mood. If not, then we aren't always in a mood. There doesn't seem to be a substantive problem here.

A more interesting conclusion is one that I've already discussed somewhat but which will be helpful to bring up again: changes in beliefs and desires *cannot* affect one's mood. This seems quite odd. Surely, if a depressed person changes enough of his beliefs he will be less depressed! However, upon examination, this consequence need not seem so problematic. All Lombard is asserting is that an *M*-state never changes as a *direct result* of practical reasoning. As he puts it "changing *M*-states is like falling asleep: one may do all one can to make it likely that one will fall asleep or that one's mood will change, but one cannot consciously or unconsciously *reason* oneself to sleep or *reason* oneself out of a mood" (Lormand 401).

This seems intuitively plausible. For example, let's say you're jealous of a friend of yours because she did better than you on an exam. If you learn that she didn't do better than you, your jealously immediate dissipates. But in the case of mood (say, being irritable) there does not seem to be any obvious new belief or desire you could acquire that would immediately dissipate the mood in a similar fashion. Whether or not a mood can change quickly or not is a matter of phenomenological debate, but it does seem clear that even if a mood can shift completely on hearing a new piece of information, this piece of information would not change one's mood in virtue of its rational relation to any beliefs and desires that make up the mood because a mood is not constituted by any beliefs or desires. In other words, if a reason changed a mood, it would not be in virtue of its being a reason that it would necessarily do so. (Unlike the jealously example, in which learning that your friend scored lower than you *necessarily* changes your emotion, it would be impossible to be jealous of your friend scoring higher than you on a test while believing that your friend did not score higher than you.) This is because moods have no *particular and specific* intentional objects (or in Lormand's account, no intentional objects at all). Except for rejecting beliefs and desires, Lormand's account leaves open the issue of what causes changes in *M*-states. As he writes, "not much can be asserted prior

to experimentation" (Lormand 400).

Nonetheless, other mood theorists have criticized Lormand's account. In "Folk, Functional, and Neurochemical Aspects of Mood," P.E. Griffiths points to a number of problems with the notion that moods are variables that inhibit or facilitate the activation of various classes of intentional states. First, he notes that moods do not only affect intentional action. They have physiological effects (Lormand only spoke of physiological states causing moods, and gave no account of how the reverse occurs), as well as direct effects on behavior; for example, in monkeys, depression is associated with huddling, rocking, and self-gripping (Griffiths 18). Many moods also affect all sorts of mental processes and even reflexes. Such physiological and behavioral effects cannot be accounted for if moods are only filters of intentional states.

Lormand's account has also been accused of circularity by Laura Sizer in "A Computational Theory of Mood." Sizer notes that Lormand does not given an account of the sets of intentional states that are characteristic of each mood. Without an account of such sets, Lormand's account seems to implicitly define a set by reference to the mood that it is supposed to explain. For example, what is it that unites all the latent intentional states activated by the mood "elation"? Well, each of these beliefs evaluates things as wonderful. But what does "evaluates things as wonderful" mean other than being a belief associated with the mood "elation." Unless the "wonderfulness" of a belief has a different criterion of evaluation, Lormand's account is circular.

Mood As a Higher-Order Functional State: P.E. Griffiths

Given the problems with Lormand's account, it's clear that a different functional description of mood is necessary. P.E. Griffiths proposes a more compelling thesis that a mood is a "higher-order functional state." To understand the meaning of this term, it may briefly help to rehash the meaning of "functional state," and compare two similar-sounding but conceptually distinct notions of "level of functional state" and "order of functional state."

A functional state is "one of a set of states of a system defined holistically by specifying the relations of those states to inputs and outputs of the system, and to one another" (Griffiths 20). When we think about the mind as a functional system, we are talking about mental states as states that cause certain types of behavior when given (sensory) inputs in conjunction with certain other states (other functional states, things like beliefs, desires, etc.). That the system is defined holistically means that functional states must be defined relative to each other. For example, one might define the functional state "desire to eat ice cream" as the tendency to eat ice cream when it is put in front of you, provided you have a belief that there is ice cream in front of you (this belief is another functional state).

Griffiths wishes to distinguish the concept of "higher-order functional state" from "higher-level functional state." The notion of a "higher-level functional state" derives from the idea that rather than there being just a level of physical description (referring to physical matter) and one level of functional description, there can be a wide variety of levels of functional description. Each level is defined by its own vocabulary, and the criterion of distinctness of functional levels is multiple realizability. Griffiths puts this more explicitly: "If a description in vocabulary V1 is realizable by more than one description in another vocabulary V2, then the two descriptions are at distinct levels, and V1 is at a higher functional level than V2" (Griffiths 20). In other words, if one functional vocabulary can be rendered in another functional vocabulary in multiple ways, then the first vocabulary is at a higher functional level than the second.

A higher-*order* functional state is not the same as a higher-level functional state. Different orders of functional description can occur at one level. Griffiths explains the notion of a hierarchy of functional states in the following passage:

> Suppose a system is capable of realizing, at different times, a number of functional descriptions F1-Fn, all at one functional level. Suppose further that the system realizes another functional description which gives as its outputs the system realizing one or other of F1-Fn. Then we have the basis for a hierar-

> chy of functional states. (Griffiths 21)

Think of a computer. When I turn on my computer, it realizes a certain functional description. When I start my word-processing program, I provide an input to the computer whose output is a new set of functional descriptions for the computer (the descriptions within the program). The easiest way to describe the computer is by describing it as a set of programs, which are produced as the outputs of a further program (the operating system). One could treat the computer as one big program, but treating it as a variety of programs is much easier and more useful.

In Griffiths's view, a person is also a system capable of realizing multiple functional descriptions. One set of descriptions applies to our normal functioning. There will be other sets that apply when an individual is in a different mood. When in a mood, the same inputs might go through many of the (liberally construed) same states,[2] but the output will be different. The mood modifies the function description of the same beliefs or desires. While it may be possible to describe the individual as a single functional system (where each same "belief" has numerous functional states corresponding to different possible moods), this is analogous to describing the computer as a whole functional apparatus rather than as containing individual programs. It is much easier to describe the human being as a system in which functional descriptions are nested within other functional descriptions. Griffiths asserts, "to understand mood and trait phenomena is to understand one or more higher-order functional descriptions of persons. These higher-order descriptions relate various kinds of inputs and outputs consisting of the per-

[2]The states are the "same" in a similar way that different people have the "same" mental state. For example, my mental state "angry about the 2004 election" is said to be the "same" as my roommate's thought "angry about the 2004 election" despite the fact that we don't have identical holistic systems of states and therefore the states cannot be exactly identical. Similarly, we can say that the "same" beliefs, desires, or other states lead to different outputs when in different moods. For example, take my belief "I am writing a paper." In an anxious mood, this belief may have one functional description (possibly it is linked to other mental states like frustration, or behavior like procrastination), but in a happy mood, the "same" belief may have another functional description (maybe it is linked to other mental states like curiosity). Here, the functional states of the "same" beliefs are not identical.

son fulfilling one of the various lower-order functional descriptions" (Griffiths 21). Understanding a mood is like understanding a computer program. Griffiths summarizes his view as the following: "Moods and traits are states with certain highly salient properties for the purposes of our psychological theories. These properties cause global changes in our propensities to occupy certain other states, and to respond in certain ways to certain stimuli" (Griffiths 23).

There are advantages to Griffiths's theory over both Lormand's view and other theories, such as the Dispositional account. First, Griffiths's theory allows moods to affect behavior both directly and through the mediation of emotions. This helps explain certain patterns of behavior that are regularly observed with mood that do not seem to require any intentional content. For example, depressed monkeys usually rock and hug themselves. It is implausible to think these types of behaviors are mediated by intentional emotional states. Second, this view does not limit moods to only predisposing one to certain "appropriate" behaviors. Moods can have a wide variety of effects. This brings the theory in line with current psychological research. Third, this theory enables emotions and other states to affect moods, since, unlike with dispositions, lower-order functional states can affect higher-order states. For example, there is an input to the word processor that takes the computer back to its normal state. What makes one order higher than another isn't causal asymmetry, but rather "in the context of some theory, it serves the interests of theory construction to introduce a hierarchical structure amongst certain properties, in virtue of the high salience within the theory of certain dispositional relations" (Griffiths 23).

Although Griffith's theory of moods as higher-order functional states sounds quite abstract, it coincides nicely with our folk psychological understanding of mood. An example will illustrate. Let's say a friend of mine is applying to law school and is nervous about getting in. He has wanted to be a lawyer for years, and he's afraid that if he doesn't get in now, he might never be able to go. But when he gets his offer of admission from his first choice school, he's not that excited. How do I explain this behavior? I could suppose that his desire to go to law school has changed. However,

that's incredibly unlikely, given that he has had this desire for a long period of time. Or maybe his belief about not being able to go if he didn't get in now has changed. But yet again, this has been a firmly held belief, and it's unlikely to have changed quickly. A more likely explanation is that he is depressed. But what does it mean to be depressed? For Griffith, depression does not lead to a change in one's beliefs, but to a change in the functional description of the same beliefs and desires[3]: "Rather than altering otherwise well grounded belief and desire attributions, we can posit this higher-order state which alters the transition probabilities between those states, and the input/output relations we need to account for" (Griffiths 24). So, while the "same" beliefs and desires remain, their output behaviors and relations to other mental states change.

Explaining the same scenario according to Lormand's theory yields a less plausible account. According to that view, my friend's depression is inhibiting the activation of certain latent intentional states while facilitating the activation of others. Presumably, the depression is inhibiting the desires to go to law school and activating the set of latent states associated with negative beliefs about law school (maybe that it will be difficult and expensive, etc.). However, this does not seem to be a plausible account of what is going on. My friend is quite consciously aware of his desire to go to law school (I keep reminding him!) and is not focusing on the negative aspects of law school (I keep accentuating the positive to him). Nevertheless, he is unhappy *in spite of* the fact that he desires to go to law school. The mood does not change the intentional states that are brought to consciousness. Rather, the mood modifies the function of the intentional states themselves.

Qualia Problem?

We have been speaking about moods as dispositions for emotions, emotions with "general" intentional content, regulators of latent intentional content,

[3]Again, to avoid confusion, it's helpful to clarify that the "same beliefs and desires" is used loosely here. Of course, these beliefs and desires are not exactly the same, since they have a different functional description.

and higher-order functional states. But, in this entire paper, we have never even discussed "what it's like" to be angry, depressed, or elated. Is this a problem? In relation to emotion, William James famously said that emotion *is* the feeling of bodily changes (James 449). Are we missing out on the essence of mood if we do not discuss the bodily sensations that accompany it?

Griffiths does not think so. He asserts that "[i]t seems that a person can be in a mood without experiencing any qualia. Someone can become irritable without realizing it, and only realize they are irritable when their friends reproach them for it" (Griffiths 28). This is probably justified. Many moods are not "conscious," so conscious feeling is not a necessary condition for some moods. However, I'm not sure if one could say that all moods could be unaccompanied by conscious feelings of some sort. It's hard to imagine someone feeling ecstatic without being aware of it. The question is how integral the qualitative experience is to the explanation of the state. If the qualitative experience merely "goes along for the ride," then it's safe to say that it can be ignored.

Another issue regarding qualia is the relation between our mood and the qualitative aspect of our perceptions. There seems to be a way in which our moods structure the qualitative aspect of our perception itself, beyond any change in belief or desire. Wittgenstein points to this in a passage from the *Remarks on the Philosophy of Psychology*: "One tie-up between moods and sense-impressions is that we use the concepts of mood to describe sense-impressions and images. We say of a musical theme, or a landscape, that it is sad, cheerful, etc." (Wittgenstein 926). What is it to say that a painting is cheerful? Is this just a metaphorical use of a mood term? Or is there something more primordial about the relationship between the qualitative aspect of our perceptions and our mood? These are difficult phenomenological issues and are beyond the scope of this paper. However, it is good to keep in mind that there may be an important aspect of mood that the functional analysis by its very nature cannot capture.

Why Functional Analysis Is On The Right Track

Despite some problems with the functionalist approach to the problem of mood, there is reason to believe that analyzing moods as functional states is more effective than attempting to define moods relative to their representational content. Theories such as the Generalization Theory that try to characterize mood by their content end up with circularity problems because there is frequently no way to group the various sets of content-full states without appealing to the moods themselves. Moreover, phenomenologically, it seems that moods bias our cognition in particular ways rather than force us into having certain intentional states. As Sizer argues, "what characterizes mood is *how* we think, not *what* we think. It suggests that what identifies and explains mood is not any set of intentional, representative states, but the processes that produce, or effect, the production of these states" (Sizer 754). Sizer uses an example of a defective camera that produces a collection of blurry photographs. When you develop the pictures, you see that something is wrong. From the photos themselves, you can tell something is amiss, but you cannot tell *why* it is amiss. To *explain* the blurriness, you have to look inside the camera; that is, a property of the underlying mechanism explains the blurriness.

Similarly, Sizer argues, philosophical theories that try to understand mood by looking at representational content are looking in the wrong place. Moods are underlying mechanisms: "anxious and joyful thoughts about work, the bath mat, the neighbor have nothing in common except that the processes which generated them were biased in a certain way" (Sizer 754). Looking for the definition of a mood by comparing its representational content in different contexts is "attempting the equivalent of analyzing hundreds of blurry photographs in order to explain the quality they all share" (Sizer 754). The power of the functional analysis of mood rests in its ability to analyze the "how" of our moods rather than solely the "what."

Works Cited

Ekman, Paul. "Moods, Emotions, Traits." *The Nature of Emotion*. Ed. Paul Ekman and Richard Davidson. New York: Oxford University Press, 1994.

Griffiths, P.E. "Folk, Functional, and Neurochemical Aspects of Mood." *Philosophical Psychology* 2 (1989): 17-33.

James, William. *Principles of Psychology*. New York: Henry Holt, 1890.

Lazarus, Richard. "The Stable and the Unstable in Emotion." *The Nature of Emotion*. ed. Paul Ekman and Richard Davidson. New York: Oxford University Press, 1994.

Lormand, Eric. "Toward a Theory of Moods." *Philosophical Studies* 47 (1985): 385-409.

Sizer, Laura. "Towards a Computational Theory of Mood." *Brit. J. Phil. Sci.* 51 (2000): 743-769.

Wittgenstein, Ludwig. *Remarks on the Philosophy of Psychology*. Oxford: Basil Blackwell, 1980.

Rigorism and Formalism
Deciding on the Content of Universal Law

Graham Leach-Krouse
Amherst College

The Universal Law formulation of Kant's categorical imperative is often troubling for ethicists, as different interpretations can yield significantly different moral results in individual cases. This paper addresses two of the most important contemporary interpretations of Kant's Universal Law formulation, namely that of Christine Korsgaard and Onora O'Neill. In the paper, Leach-Krouse argues that no satisfactory compromise between these two interpretations can be found. Specifically, the papers argues that Korsgaard's account avoids rigorism—the over broad application of moral rules—by interpreting the Universal Law Formula as a guide to particular actions, and that O'Neill's account avoids formalism—the failure to prohibit immoral acts—by interpreting the formula as a source of overarching rules for conduct. Leach-Krouse argues that by examining morally indifferent actions that are similar in certain ways to morally unacceptable actions, we see that there can be no acceptable midpoint of maxim-generality that avoids both formalism and rigorism.

A number of potential problems with Kant's ethical system are rooted in the Formula of Universal Law (shortened to "FUL" in what follows), a formula which is foundational not only to Kant's moral dictates, but also to his closely associated ideas on free will and the practical capacity of reason. Different broad categories of difficulty with the FUL can be distinguished. It is not entirely apparent, for example, that the FUL is equivalent to the Formula of the End in Itself, which Kant generally takes to follow from it. Alternately, one could argue that Kant's concept of duty—behavior valuable for merely its own sake—does not accord with any number of popular

definitions of the good, or goodness. It may not, in fact, have much in common with the colloquial notion of duty as such. Although the FUL draws some of its plausibility from our pre-theoretical admiration for duty, it may abstract away from exactly the phenomenon it hopes to formalize.

But the most direct criticisms of the FUL, and the ones with which I will concern myself here, allege that the FUL is faulty in terms of content. It may forbid acts that we would almost certainly regard as ethical, or refuse to ban acts that seem unacceptable. Of course, a certain amount of counter-intuitiveness is forgivable. A large part of Kant's appeal lies in his expectation that there be substantial and consistent reasons behind our ethical judgments. If our judgments do not accord properly with Kant, it may be that they are in some way inconsistent or in need of revision. Christine Korsgaard points out that Kant's ideas about sexual equality were ahead of his time ("an unequal marriage is not a marriage at all. Thus, marriage as practiced in most societies has not sanctified but rather degraded sexual relations" (Korsgaard 195)), despite his ideas on psychological sexual difference being somewhat less so. But even giving Kant the benefit of the doubt, there are a few ethical data that we cannot abandon, no matter how inconsistent they are with Kant's system. I think the objections that I plan to present draw from that category.

In the first following section, I will reconstruct the relevantly problematic dictates of the FUL as it is stated in *The Groundwork of the Metaphysics of Morals*. In the second section, I will address a few preliminary content complaints, along with recent revisions and interpretations of the FUL that seem to overcome them. And in the third section I will present remaining systematic difficulties with the FUL.

Definitions of Universal Law

The Formula of Universal Law gives us a remarkable amount of ethical mileage, considering that it is only one sentence long. Kant, and allegedly reason itself, tells us that we should "Act only on that maxim through which you can at the same time will that it should become a universal law" (Kant,

1948: 52). A slightly different formulation on the same page tells us that "Since the universality of law governing the production of effects constitutes what is properly called 'nature'... the universal imperative of duty may also run as follows: 'Act as if the maxim of your action were to become through your will a universal law of nature' " (52). And a third variation later in the *Groundwork* demands that we "Act on the maxim which can at the same time be made a universal law" (81), making it clear that we can only will as universal laws maxims which could coherently *be* universal laws.

Maxims

An exact notion of a maxim is important to understanding any of those variations of the FUL. Kant defines a maxim as "a subjective principle of action... a practical rule determined by reason in accordance with the conditions of the subject... a principle on which a subject acts" (51n). This may seem slightly unclear, but perhaps we can paraphrase as follows: a maxim is the principle, policy, or rule of thumb expressed[1] by an act. The maxim of a fraudulent debtor could be something along the lines of "Whenever I believe myself short of money, I will borrow money and promise to pay it back, though I know this will never be done" (54). Given that Kant tells us that the "ground of [the will's] self-determination is an end" (63), and that "rational nature separates itself from all other things by the fact that it sets itself an end" (82), it seems plausible to add that Kant believes purposiveness is a necessary feature of all really human behavior (as distinguished from reflex or pathological habit). We might then think of maxims as the products of ends and practical rules for obtaining ends, or "hypothetical imperatives" (39).

What it might mean for a maxim to "become a universal law" (52) seems best understood by imagining that that maxim will become a "law governing the production of effects" (52), and will dictate how everyone actually does

[1] In the way that an object's tendency to remain at rest unless acted on by an outside force might *express* a principle of inertia, for example.

act, rather than just place normative constraints on how they ought to act. If the alternative seems tempting, it might be helpful to assume "ought implies can," and take the possibility of everyone acting a certain way to be implied by the possibility that they ought to act that way. In any case, I will take it that the possibility of willing a maxim as universal law depends on the possibility of everyone actually acting according to that maxim.

What might everyone acting according to my maxim amount to? It seems like this would mean everyone behaving as I do, that is, adopting the end and accepting the hypothetical imperative that my maxim involves. What this means in practice needs some clarification. Say I am eating a sandwich. The only thing that is obvious is the result of the relevant hypothetical imperative: "If X, you should eat your sandwich." To discover more, Onora O'Neill suggests an "isolation test"—what Kant uses elsewhere (10-11) to discover if an action is motivated by duty or pleasure. If an act is not being carried out from a motive of pleasure, then if pleasure is removed, the act will not be disturbed. So, in the sandwich case, we might ask whether I would still eat if I were not hungry, if I did not enjoy sandwiches, and so on. If the act would continue, then I can conclude that those motives were no part of my end. Of course, if the act would cease, the motives in question may have been merely part of my end.[2] This, combined with the potential for error in imagining the counterfactuals involved, makes it difficult to ever be certain of our own maxims. Isolation testing does give us a hint of what the universalization of a maxim might look like, however. If the end of my maxim is to abate hunger, then the maxim, although universally held, will only be acted on in cases where hunger is present. The conditions discovered by isolation tests, which must be met in my case to transform my maxim into action, will result in action when met by others. So, if the maxim of my action is universalized, everyone will perform that action when in *relevantly similar circumstances*.

[2]What if my end is to pleasurably abate hunger? Then having either pleasure or hunger removed individually will still keep me from serving my end by eating the sandwich.

Contradiction In Conception

There are two ways in which I can fail to fulfill the FUL, but one of them in particular gives impetus to the charges of faulty content. This is what is called a "contradiction in conception." A contradiction in conception occurs when my action might "be so constituted that [its] maxim cannot even be conceived as a universal law of nature without contradiction" (57). Kant's well known dishonest loan example is a contradiction of this kind. If I made it my maxim to obtain money by disingenuously promising that I will someday return it, then the universalization of my maxim would result in a world in which my kind of behavior were the rule rather than the exception. Since loan applications would nearly always be dishonest, "everyone would laugh at promises of this kind as empty shams" (50) and acts guided by my maxim would always fail. It is not actually possible for everyone to act according to that maxim, and for that reason, according to Kant, willing the universalization of such a maxim is not possible. A maxim which results in this sort of contradiction "is opposed to strict or narrow (rigorous) duty" (57)—which is for all intents and purposes the same as "perfect" duty, and "allows no exception in the interests of inclination" (53n). By testing for contradictions in conception, we can create a whole body of moral law over maxims. Kant also describes another kind of contradiction, i.e. contradiction in the will. Contradiction in the will only gives us an idea of "wider (meritorious) duty" (57), which does allow exception in the interests of inclination. But since we are only interested here in whether the FUL permits or strictly denies moral and immoral actions, we can pass over the optional moral advice given by wider duty.

Definitions of Formalism and Rigorism

There are two ways that the contradiction in conception test might be faulty in terms of content. It may fail to prohibit certain obviously immoral acts, or it may reject acts or omissions that are obviously indifferent (or morally worthy). We can call Kant's theory "formalistic" if it fails to give suffi-

ciently strict moral instructions. If it applies rules too broadly, condemning inappropriately, we can call it "rigoristic." Both terms are sometimes used flexibly elsewhere in the literature. For example, it is not always clear if formalism means that a theory is completely empty, or just importantly empty. Similarly, we might wonder whether a rule needs to be applied in all cases, or simply where it is problematic, to be called rigoristic. I will use both terms in the broader sense; formalistic should be taken to mean "importantly empty," and rigoristic to mean "problematically strict." I will present some of the defenses against formalism first.

Formalism

Depending on what is meant by the contradiction in conception test, certain obvious crimes might slip past it. It is imaginable that I could make it my maxim to take what I want whenever I can get away with it, without ever causing a breakdown in the conditions that make the physical act of theft possible. False promises rely on a standard of trust that could be eroded by universalization, but the conditions for taking something that does not belong to you are more or less natural, rather than conventional, and for that reason far less fragile than the conditions that make, for example, loans possible. Christine Korsgaard's "practical" interpretation builds on the purposive aspect of maxims and the features of the Kantian conception of willing in order to bulwark the contradiction in conception test against this kind of problem.

Korsgaard's practical interpretation draws on Kant's specification that we must be able to act on our maxim and *at the same time* will it as universal law (52). If our end in originally willing our act is incompatible with the effects of the universalization of our maxim, then the maxim is contradictory. One can steal successfully even if theft is universally practiced. But a thief destroys another person's ownership of an object in order to secure the same kind of ownership for himself. The universalization of his maxim of theft creates a world in which no one can securely own the sort of thing the thief intends to acquire. Universalizing a maxim that aims at owner-

ship through theft creates a world in which ownership is impossible. The criminal is unable to universalize his maxim, because his maxim works by depriving others of what the thief desires, introducing an asymmetry that favors the thief. As Korsgaard reminds us "the kind of case around which [Kant's] view is framed, and which it handles best, is the temptation to make oneself an exception: selfishness, meanness, advantage-taking, and disregard for the rights of others" (Korsgaard 100-101). If she is correct, her interpretation seems to capture the spirit of the FUL quite well.

One could object here that simultaneously willing that one's ends obtain and willing that a world in which those ends do not obtain is not strictly speaking a contradiction, but rather a violation of an artificially imposed rule of consistency. If what I will is just a mental quality of mine, then I could will mutually exclusive things, just as easily as I could desire, for example, both to go to the movies and to go to the theater on a given night. A contradiction would be to desire to go to the theatre and to not desire to go to the theatre—i.e. to have and not have some desire-predicate, rather than to desire to go to the theatre, and to desire not to go to the theatre—i.e. to have two different desires. Kant's system, however, is already prepared to handle this kind of objection. To see how, we need to look at his conception of the will:

> Who wills the end, wills (so far as reason has decisive influence on his actions) also the means which are indispensably necessary and in his power. So far as willing is concerned, this proposition is analytic: for in my willing of an object as an effect there is already conceived the causality of myself as an acting cause—that is, the use of means; and from the concept of willing an end the imperative merely extracts the concept of actions necessary to this end. (Kant, 1948: 45)

Willing is to have a particular kind of picture of the world, one in which "there is already conceived the causality of myself as an acting cause" (45). If I am to be the cause of some end, I need to bring about whatever is necessary to produce that end. So for my picture of myself as the cause to be

consistent, I need to include acting on the means to the end in it: in willing the end, I will the means by definition. And similarly, I cannot will two mutually exclusive ends without contradiction. Mutually exclusive willings are contradictory because Kant's conception of the will is not amorphously introspective: it demands action.

But even under the practical interpretation of contradiction in conception, cases remain in which a maxim that involves gain at the expense of others can be willed and universalized. These cases usually utilize a maxim with highly specific relevantly similar circumstances. If the conditions that provoke action are rare enough, such a maxim will never be acted on in a way that interferes with the maxim holder. For example, a mother might make it her maxim to get some peace and quiet by killing a crying infant (which it seems, would deprive the infant of rest). Even if this were universalized, assuming the child were exceptionally loud (louder than the killer was when she was an infant), the killer would still be able to achieve her end. Other immoral acts might not deprive the victim of exactly what the criminal maxim aims at, if the aim is sufficiently specific. For example, Korsgaard admits that people who kill out of revenge or hatred are unlikely to find their end frustrated by universalization. Assuming their ends are purely murderous, and do not entail survival or security, they will not necessarily be frustrated by the violent but conceivable state of nature that universalization of their maxim implies.

O'Neill's solution to cases of this kind emphasizes that "both instrumental and brute violence undercut the agency of those whom they victimize. It is not merely that victims do not in fact will the maxims of their destroyers and coercers: They are deliberately made unable to do so" (O'Neill, 1989: 133). If we destroy someone, they never have a chance to make destruction their maxim—it is not possible for everyone to actually act according to a maxim of destruction, because someone will have to be destroyed without destroying. And if it is not possible for everyone to actually act on a maxim of destruction, then that maxim cannot be willed as universal law. Thinking about contradiction in conception this way supposes that maxims are much more general than the principles of individual actions; on O'Neill's inter-

pretation, a maxim is involved in the decision to lead a violent life, not just in the decision to use violence in any particular circumstance.

O'Neill can deal with the kind of particular intentions that give Korsgaard trouble. We should pause here to look at the background that makes this sort of solution possible. As noted in the first section, it is not always obvious what the maxim of any given act might be. The maxim of vengefully murdering a particular person might be stated "I will kill the individual Ignatz MacGillycuddy on sight if he wrongs me, to obtain revenge," and this could be universalized without much difficulty at all (even Ignatz could hold it passively, although he is unlikely to wrong himself). Likewise, "I will kill overloud children to get some rest." Both murders, however, also could express a broader policy: "I will kill to secure advantage for myself." That maxim is clearly *not* universalizable by O'Neill or Korsgaard's standards. The permissibility of destructive acts in Korsgaard's account depends on the acts having such a limited scope of intention. But it seems plausible that these acts express a far more general immoral habit. We can think of several sets of means and ends that could have motivated each murder—we can call these "potential maxims"—but only one of these can be the actual maxim. If that intuition is correct, then "The fact that we can formulate *some* universalizable surface intention for any action ... is no embarrassment to a universality test that is intended to apply to agents' maxims" (O'Neill, 1989: 87).

To discover which of the many potential maxims associated with a given act is morally relevant, O'Neill's interpretation of the FUL emphasizes that "*Maxims are those underlying principles or intentions by which we guide and control our more specific intentions*" (1989: 84). Maxims are more like character traits than results of any spur of the moment decision-making. The vast majority of the potential maxims for any action are irrelevant:

> For principles of action only need to incorporate some true description of the agent and some true description of the act and situation, whether those descriptions are vacuous or brimming with detail. But an agent's maxim in a given act must incorporate just those descriptions of the agent, the act, and the situa-

> tion upon which the doing of the act depends. (1989: 84)

In practice, this means that our actual maxims should be understood as independent of our circumstances, and often even of our desires and conscious intentions. If I hold a maxim of welcoming guests, I might serve tea in England and wine in France. The maxim only entails that I greet the guests—how I greet them is immaterial. What really distinguishes a genuine maxim from the potential ones is how we would behave if things were otherwise. If I hold a maxim of greeting guests in England with tea, then I will be at a loss in France. The unlikeliness of such an inability to generalize my fondness for English guests attests to the unlikeliness of my holding such a specific maxim.

We can get at the underlying principles of an action with isolation testing. For example if my murderousness really is only piqued by Ignatz, I would refrain from violence in other cases. If, on the other hand, I would attack a range of individuals when they wronged me, then it cannot be Ignatz "upon which the doing of the act depends" (O'Neill, 1989: 84). Instead, the maxim must be a more general policy of revenge. Indeed, it does seem to border on absurdity to suppose that someone could be fundamentally the sort of person who would kill Ignatz, but only Ignatz.

Broad rules of conduct can prohibit the kinds of acts that slipped by Korsgaard's contradiction in conception test: "precisely because the Categorical Imperative formulates a universality test that applies to maxims, and not just to any intention, it is not rebutted by the fact that relevantly specific intentions often can be universalized without conceptual contradiction" (1989: 97). It seems that thinking of maxims as underlying broad principles can fill in most of the ethical gaps which motivate by the formalist objection.

Rigorism

One of the simplest objections to the FUL is that it seems as if many obviously morally neutral maxims cannot be universalized—that is, the problem of rigorism. For example, going to the park to enjoy a sunny day encounters

serious difficulties. If everyone went to the park, it would be closed because of dangerous overcrowding; universalizing the maxim makes it impossible for someone to act on the maxim. It seems as if the feasibility of this maxim does depend on the (moderate) exceptionality of my action. Of course, it is possible that I am misstating my maxim. If we are following O'Neill in conceiving of maxims as underlying principles, then my maxim might just be "If it is pleasant out and you have free time, go outside and enjoy yourself," which can be universalized without difficulty.

An isolation test here can help us distinguish between the two potential maxims and determine the kind of behavior actually specified by my maxim. We can test by considering whether I would still act if I were prevented from reaching my goal in a particular way. If I am prevented from enjoying myself at the park, do I fail to act? Or do I find some other way of fulfilling my maxim? If I still act on my maxim when the element targeted by the isolation test is removed, then that element is not part of my maxim. If I still go out to enjoy the day when the park is closed, then going to the park was not in fact the behavior specified by my end; rather, I was acting according to a more general command that happened to lead me to the park.

This shows that the action of going to the park, at least, is potentially innocent. But what is so wrong with holding a maxim of going to the park, rather than a general maxim of going outside? Is this an instance of rigorism? I do not think it is. Though it is not obvious, by holding a maxim like "To enjoy the sun, I will go to the park," I may be taking advantage of those who do not arrive at the park in time to enjoy it. We can make the moral aspect of the maxim more obvious by raising the stakes. Consider a parallel case of limited resources: a sinking ship with too few lifeboats for the passengers. The maxim "I will board the lifeboats to save my life" cannot be universalized, for the same reason that the park example cannot be—it is not possible for everyone to act on such a maxim, since the lifeboats will rapidly overcrowd. The comparative disadvantage of the weaker passengers becomes means to my end—exactly the situation O'Neill thinks that the FUL tells us generally to avoid. On the other hand, getting an agreed

upon group to the lifeboats (or to an undersized park) *is* universalizable.[3] Drawing straws is likewise acceptable, as is, interestingly, the crew's honorably going down with the ship.

However, Kant brought a worse cluster of rigorism accusations down on himself with the late essay *On a Supposed Right to Lie from Altruistic Motives*. Benjamin Constant in *On Political Reactions* suggested that it would be acceptable to lie about the location of one's friend, when a person looking to kill him arrives at your front door (Kant, 1996: 611). Kant replied, bizarrely, that not only was it unacceptable to lie, but the liar may be held "legally as well as ethically responsible for the consequences" (Korsgaard 133) if a fantastic string of coincidences connected the lie to the eventual success of the murderer.

It is fair to say there is a problem here. Is it of a kind that O'Neill can accommodate? What would it demonstrate about our underlying principles if we lied to save the life of our friend? It is not obviously implausible that a really serious Kantian would lie to save a life. But neither is it impossible, as Kant himself demonstrates, that he would only save lives as long as it did not involve lying. There seems to be a conflict between two broad underlying principles, one of lifesaving and one of honesty. Very general rules can encroach on each other, and when they do, O'Neill's prohibition against "relevantly specific intentions" (O'Neill, 1989: 97) seems problematic. If I save my friend's life, it does seem to be a one-time action, in a violent situation. If I ignore the lifesaving rule in this case only in order to stay honest, I will still probably obey it when it is uncontroversial.

In early writings, O'Neill admits that it may be possible to universalize either kind of relevantly specific underlying principle (O'Neill, 1975: 135)—"lie only to save lives" or "permit deaths only when honesty requires." There is no clear decision procedure for telling which exception is the moral one, except perhaps an unhappy compromise implied in *On the*

[3]Even for members of that group, as long as their support for the policy doesn't depend on their membership in the group—if isolation testing reveals that they would not save the group if they weren't a member, then their maxim is probably more like "I will board the lifeboats to save my group," which isn't universalizable for the same reasons that "I will board the lifeboats to save myself" isn't.

Supposed Right to Lie. An unreformed maxim of honesty requires unethical actions. But if we evaluate principles at the level of generality that O'Neill suggests, we find that lying to the murderer is forbidden. To compound the trouble, it seems that not helping your friend is permissible in some cases, while lying is never permissible. We have a perfect duty not to lie, since lying to get what you want cannot be universalized. But it is unclear that we have a perfect duty to always keep others from coming to harm. A person might "have no wish to contribute [to another's] support in distress. [And] admittedly if such an attitude were a universal law of nature, mankind could get on perfectly well" (Kant, 1948: 56). If they are separate duties, with no relevant specificity, they conflict. And if we justify overruling perfect duties by appeal to imperfect duties in this case, we open the door to a swarm of problems. Although we *can* formulate a relevantly specific maxim that will get us through *On the Supposed Right to Lie*, O'Neill can't tell us why we *should*.

Korsgaard, understandably, wants a stronger solution, and argues that Kant is mistaken in his application of the contradiction in conception test here. If the murderer is trying to trick you, Korsgaard argues, and he assumes that you don't know he's a murderer when he opens the door (which seems reasonable), then

> the lie will be efficacious even if universally practiced. But the reason it will be efficacious is rather odd: it is because the murderer supposes that you do not know what circumstances you are in—that is, that you do not know that you are addressing a murderer—and so does not conclude from the fact that people in those circumstances always lie that you will lie. (Korsgaard 136)

Again, we need to step back and look at the background that makes this kind of solution possible. Korsgaard picks out maxims at a relatively specific level for testing. Because "willing is regarding yourself as the cause of the end in question" (Korsgaard 94), what we will is linked quite closely to what we regard as our ends in acting. Our maxims, in Korsgaard's picture, are our self-attributed intentions and goals. This is quite different from

the kind of underlying principle of conduct that O'Neill associates with a maxim. Trivial ends like going to the park or picking a certain food at the buffet may be below the level of intention, which avoids the kind of limited resource dilemma we saw above. But unlike O'Neill, Korsgaard can appeal to specific intentions to get out of specific bad situations.

There does seem to be a sense in which we intend to hold the broad underlying principles our actions express just as much as we intend to carry out individual actions. This seems to imply that there could be more than one valid way to describe the maxim of an action, since we can appeal to both the particular intention and the deeper underlying intention. Korsgaard deals with the overabundance of maxim descriptions by breaking them down hierarchically. Specific maxims after all "must be adopted for a reason…[so] there are reasons for having purposes, which are again expressed in maxims" (Korsgaard 58). Our underlying principles form the ground floor that supports the adoption of the more specific intentions involved in acting.

This kind of federal decision-making lets us break the rules, in a way. When we adopt a general policy like "To protect life, avoid violence," we only need to check if the adoption of a similar maxim by everyone would be contradictory. Once the policy is in place, it can issue situation specific maxims such as "To avoid violence, lie to murderers," and these can be evaluated for consistency with the FUL separately. It is still true that "[i]n most cases lying falls squarely into the category of the sort of action Kant considers wrong: actions whose efficacy depends upon the fact that most people do not engage in them" (Korsgaard 136). But it is possible for us to have moral reasons that motivate us to lie in exceptional cases. And both the maxim of that lie, given the special circumstances, and every part of the chain of command grounded in something like "Preserve human life for the sake of duty" might pass the contradiction in conception test.[4]

[4]It's interesting to note that Korsgaard concludes elsewhere that the Formula of the End in Itself, which O'Neill holds, does in fact demand that we are honest with the murderer. But unlike O'Neill, Korsgaard does not take the Formula of the End in Itself to follow from the FUL.

Differences of Scope

By breaking down maxims, we change the scope of relevantly similar situations. Korsgaard's approach licenses us not even to lie to all murderers, but only to murderers who are lying to us. But O'Neill lacks an account of how we might have relatively specific maxims to deal with exceptional situations. In her model, a line is drawn straight from action to underlying principle, with no door for specificity and universalizable exception along the way. An account that contains this kind of conflict, rather than sidestepping it as Korsgaard does, tends towards rigorism in the *Right to Lie* case because of perfect duty. On the other hand, that same feature of O'Neill's account that involves her in rigorism shields against the sort of over-specific intention or "diseased purpose" (Korsgaard 100) that is problematic for Korsgaard.

Tension Between Rigorism and Formalism

O'Neill and Korsgaard's points of disagreement seem to center on Korsgaard's tendency towards formalism and O'Neill's (and Kant's) tendency towards rigorism. This suggests an underlying problem that neither philosopher fully addresses. Modifications to make Kant's system less rigorous or less formal pull in opposite directions. It is difficult to see how one problem can be corrected without making the other worse.

The Technicality of the FUL

The trouble might be traced to the fact that a universalization test is an indirect route to our moral data. The FUL seems to be an attempt at a moral doctrine with "an origin completely and entirely a priori" (Kant, 1948: 60), one that a person can apply without knowing anything beforehand about what is moral and what is not. If it succeeds, it avoids any sort of subjectivism in morality. But the right kind of definition would need to evaluate acts for morality on the basis of their non-moral features. What we need to be able to say to make the FUL work as intended is that *all* maxims

with a certain kind of relation to the world are immoral. We could say that something like

(1) When X I will A to B

(2) If everyone As When X, then C

(3) C prevents (A or B)

implies that "When X I will A to B" is an immoral maxim. But with the right kind of phrasing, we can sneak obviously immoral actions past this kind of test, or find moral actions that it condemns. Being very specific about what counts as a maxim (for example, by stipulating that it must be a general underlying principle, or alternately a conscious intention) lets us limit the range of possible inputs in (1). But it still gives us a purely technical result, something like

(1) When X I will A to B

(2) A, B, and, X are of particular limited kinds Q, R, and T, respectively

(3) If everyone As when X, then C

(4) C prevents (A or B)

And we still need to be able to say that *every* principle of this form is forbidden, and that actions following from principles not of this form are acceptable. Working with examples allows commentators to dodge the bullet of explicitly considering this implausibly general claim. But if there are kinds Q, R, and T, I do not think they have been precisely located. Variations along a continuum of generality do not seem to have a stopping point that separates moral maxims from immoral ones—if we set our values to rule out all immorality, we have gone too far, and if we permit everything that we think is acceptable, we tend to permit problematic maxims as well.

Parallel Cases

We can bring this out by suggesting a morally neutral situation that runs closely parallel to Kant's example of an immoral maxim involved in borrowing dishonestly. What makes taking the loan immoral is that it relies on the ignorance of the lender. If everyone did the same, or if I announced my intentions, the action I currently intend would be senseless. Imagine that I aim to please my friend by surprising her with a meal I have cooked. Since I am not a very good cook, my goal of beneficence depends entirely on the unexpectedness of the meal. If this were the standard practice for pleasing one's friend—that is, if this were universally adopted in all relevantly similar circumstances—then a meal would never be a pleasant surprise. Presumably, "everyone would laugh at [surprises] of this kind as empty shams" (Kant, 1948: 55). While this certainly is not as bad as, for example, making a false promise, it seems as if it would have the same sort of consequences if universalized.

Perhaps this is because my maxim behind the meal is being stated incorrectly. It might be modified by isolation tests, to discover a deeper motivation behind my actions. If I could not please my friend by preparing a meal, maybe I would give her a gift, or help her with some work, demonstrating that I act on a maxim such as "To make the world a better place, be kind to your friends." But then, couldn't the false promising example be modified by an isolation test? If the person could not get money by applying for a false loan, they might rob a bank or get a job. It is rare to intend to deceive for the sake of deception; generally our purposes only circumstantially entail deception. A maxim of money acquisition could pass the contradiction in conception test, assuming some variation in its application, and could conceivably underlie criminal actions. Trying to acquire money criminally is less likely to be permissible, and is certainly more likely to motivate a false loan. But the two intentions have a large number of unacceptable applications in common. One could specify that any sort of intention of profit that could lead to a bank robbery would be ununiversalizable, by for example cashing out the maxim involved into a collection of more specific ones: to earn money, work hard; to earn money when desperate, apply for charity;

and so on, excluding means that are not universalizable. But if you apply the same analysis elsewhere, you outlaw the kind of altruism that might result in a surprise meal or a trip to the park. Distinguishing between malignantly and helpfully general maxims, cashing out one and not the other, would require prior moral knowledge from outside the Formula of Universal law. Korsgaard does not directly address this problem, but she comes close when she rejects a challenge from Hegel, presented by F. H. Bradley:

> 'steal property' is a contradiction, for it destroys property, and with it the possibility of theft... the essence of morality was a similar contradiction... morality is... as inconsistent as theft. 'Succor the poor' both negates and presupposes (hence posits) poverty: as Blake comically says:
>
> > *Pity would be no more,*
> > *If we did not make somebody poor.*
> > (Korsgaard 87)

The idea here is that in the same way a maxim like "To secure property I will steal" undermines itself, by making secure property impossible, a maxim like "To help the poor, I will give them money" undermines itself by eliminating the poor. Korsgaard's response is that "Succor the poor" is a poorly formed maxim. "The world of the universalized maxim only contradicts one's will if it thwarts one's purpose. A world without poverty does not contradict [one's purpose in giving to the poor] but rather satisfies it another (better) way, and no contradiction arises" (Korsgaard 95). But if the details of our intended purpose in giving to the poor can save our maxim from contradiction, it is hard to avoid saying that if a thief's intention would not be thwarted by universalization of theft and the subsequent destruction of property,[5] then the FUL has no claim on him.

Rather than focusing on intentions, we could go further, and ask about the kind of underlying principle that concealing my plan to cook a meal for

[5] Imagine he is an anarchist who steals treasured heirlooms specifically in order to destroy the institution of property. Or perhaps he steals painkillers from the elderly and intends to abuse them as soon as they are in his hands.

my friend and concealing my plan to not repay a loan, respectively embody. One, we might conclude, follows from a fundamental benevolence. The other, we could find, follows from a fundamental willingness to lie. But can benevolence justify a lie, when one is at the same time fundamentally unwilling to lie? If it can overrule a perfect duty, then it can justify nearly anything, and Kant's theory collapses. If it cannot, then one of the two maxims (benevolence or honesty) needs to introduce some element of relevant specificity to be compatible with the other. But then what is to prevent appeal to similar relevant specificities in cases of bad behavior?

O'Neill argues that our maxims simply are what they are, and we cannot just appeal to surface intentions. Although we can concoct psychologically improbable scenarios, "a certain proposed act may be one that could not (in the actual circumstances) be performed by somebody whose maxim was a maxim of duty; such an act would be forbidden" (1989: 135). This helps somewhat—a wealthy person obviously is not (in actual circumstances), stealing a loaf of bread on a maxim of stealing only to feed his children, and it is doubtful that anyone ever has an underlying principle of harming only a specific individual. But not all cases of explanation by exception are so clear-cut. A specific universalizable maxim such as "I will steal from Woolworth's once yearly if I can get away with it" may well involve an "organization of other, less fundamental intentions [that reveal] it really was subject to those restriction" (1989: 97). Woolworth's aside, a thief who steals when he cannot get away with it is a pretty poor thief. We should expect their behavior to demonstrate a certain amount of caution in all cases. Talk about the relevant specificity of underlying maxims tends to be ad hoc, which is exactly the reason O'Neill rules it out. But without some degree of relevant specificity, or some degree of rigorism, we can make very little sense of a surprise meal at all.

Conclusion

As we clear cobwebs and investigate dark corners, it becomes apparent that neither O'Neill nor Korsgaard can convincingly keep morality and immoral-

ity distinct using universal law alone. Obviously there will be some give and take between our folk moral concepts and our formalized moral theories. But while the FUL can be phrased in such a way that it gives definitive moral laws, it never seems to fully accommodate our intuitions. Interpreters end up playing tug of war, elevating certain folk intuitions above others in trying to ventriloquize exactly what the FUL says. Working upwards from examples and moral data can be misleading. Thinking about some problems or types of problems may suggest apparently appropriate variations on the formula. But just filling in what we need to address some limited set that needs addressing does little to anticipate permutations that are not part of that set. Once the Q, R, and T are decided on, it is not difficult to come up with variations, and often plausible ones, designed to elude the test. What is needed is a reason *why* every maxim produced by a certain variation is forbidden—probably a reason found outside of Kant, since Kant's interpreters, including himself, have had such difficulty finding the reason for one level of specificity as opposed to another within his writings.

Works Cited

Kant, Immanuel. *Groundwork of the Metaphysics of Morals*. trans. H.J. Paton. New York: Routledge, 1948.

———. *Practical Philosophy*. ed. Mary Gregor. Cambridge: Cambridge University Press, 1996.

Korsgaard, Christine. *Creating the Kingdom of Ends*. Cambridge: Cambridge University Press, 1996.

O'Neill, Onora. *Acting on Principle: An Essay in Kantian Ethics*. London: Columbia University Press, 1975.

———. *Constructions of Reason*. Cambridge: Cambridge University Press, 1989.

Reheating the Ball of Wax

Descartes on Perception and What It Is To Be A Thing

Enoch Lambert
Cornell University

René Descartes's Second Meditation famously begins with the philosopher's musings on a melting ball of wax. Within this simple narrative, Descartes and generations of philosophers after him found passageways to many of modern philosophy's most troubling metaphysical and epistemological problems. In this paper, Lambert summarizes Descartes's theoretical treatment of the wax, and then argues that this treatment is inadequate. By unpacking the implications of the wax experiment further, Lambert argues that a large number of distinct and troubling questions are implied, but left unresolved, by Descartes. Moreover, the implications that Descartes hopes to draw from the argument, Lambert argues, are unfounded. After critiquing Descartes's approach to the wax argument, the paper takes up one particularly provoking issue—the nature of perceptual content. Here, Lambert argues in favor of the phenomenological approach to this problem, and offers his own contribution to this brand of response.

The Second Meditation contains Descartes's thought-provoking reflections on a ball of wax. His fireside experiment opens up many compelling epistemological and metaphysical questions, some of which (though not all, as I shall argue) are articulated, addressed, and eventually answered by Descartes. In this paper, I shall try to distinguish and articulate the various questions implied by Descartes's discussion of the ball of wax. I will also argue against the adequacy of the philosopher's answers. Because, as I shall argue, Descartes neither appreciates nor articulates the differences between the different questions his discussion implies, *and* because he does

not adequately answer the ones he does address, it seems that those questions remain relevant to philosophical inquiry today and deserve more attention. Finally, though I will not be able to propose my own answers to all the questions raised in the confines of this paper, I will end by suggesting some different approaches to them, including a more sustained discussion of a recent relevant debate.

Through his reflections Descartes claims to show that 1) the essence of the ball of wax does not reside in any of its sensible qualities, and that 2) he thus perceives the wax through the faculty of judgment in his mind, or pure "intellection," alone. It is not entirely clear how much philosophical work Descartes wants this discussion of the wax and its results to do for him. Does it imply (since Descartes never *explicitly* makes this claim) that *all* material bodies are perceived with the mind alone, and thus that *all* of our ideas are in some sense innate? Or is the discussion merely a rhetorical strategy for further convincing the reader of the philosophical uselessness of the information provided by the senses and imagination? This latter possibility is suggested by the introduction to the discussion, in which Descartes admits the difficulty of ignoring the senses and so allows his mind once more to consider the objects of sense reports. Once the discussion is finished, and the senses have supposedly been shown to be inadequate, Descartes immediately begins Meditation Three by announcing that he will shut down all his senses and regard all corporeal images as "false and worthless." However, even if this latter strategy is all Descartes intends, he does so at the expense of introducing a challenging philosophical problem (i.e. just what *is* a thing?) along with answers with epistemological and ontological commitments. With the ball of wax discussion, Descartes challenges our commonsense notions of what a thing is and how we can know it. Are his answers adequate? In the next section I lay out Descartes's arguments and begin to evaluate them.

Descartes's Discussion[1]

Descartes begins by saying that he needs to investigate a particular body, for this will be much clearer and more distinct to his mind than "bodies in general" (Descartes 32b). So he asks the reader to consider a piece of wax that he has in his hand. He indicates several of its secondary qualities that are available to the five senses: its honey taste, lingering scent of flowers, color, its "easy touch," and the sound it emits when rapped upon. Besides making sure he covers all the senses, Descartes also mixes in primary qualities he observes (e.g. size and shape) to ensure that *all* sensible qualities are put under equal consideration.[2] Having set up this picture, Descartes plans to show us that we still identify the same piece of wax even though all of its enumerable qualities have changed. He tells us he brings the wax near a fire and as he does so, it loses its taste; the smell, color, size, and shape all change; it no longer emits a sound when rapped upon; and it becomes hot (presumably eliminating its "easy touch"). After every quality has changed, Descartes asks if the same wax remains. He answers: "I must confess that it does; no one denies it; no one thinks otherwise" (32b). But how could we possibly know this, if all qualities of the wax have changed? If I were shown an object and then, later, another one with all different qualities, I would never judge that the two were the same. What is different in this case? The problem raises both the epistemological question of how we know and identify a thing and the ontological question of just what a thing is.

Descartes's answer to the epistemological question of how we know that the ball of wax remains the same ball of wax through all apparent qualitative changes is that we "perceive it through the mind alone" (33a). But he is at pains to show that we do not perceive it through mental images. The

[1]The text cited has two columns to a page. I will refer to the page and column throughout. The left hand column will be referred to as "a" and the right hand one as "b." Thus, e.g., a quote from the left hand column of page 49 will be cited as "49a."

[2]While Descartes may not use Locke's later terms of "primary" and "secondary" to describe the distinction, it is clear that he does make the distinction both before and after the ball of wax discussion. See 31a, 45a, and especially 49a.

aspect of the mind that represents images of physical bodies to us, Descartes calls the "imagination," and says that this cannot be what perceives the ball of wax. Why is this? He says that although he can grasp that the wax is capable of innumerable mutations in its shape, he can only represent a small number of those changes to himself by means of the imagination. Later, in Meditation Six, he makes the point by arguing that while we cannot imagine a chiliagon (a thousand sided figure), we can certainly conceive of one (presumably because we have the relevant concepts) (48a-b). In the discussion, this point serves as the basis for Descartes's distinction between different faculties of the mind, namely the "imagination" and the "faculty of judgment," or what he later calls "pure intellection" ("pure" meaning without any images of physical qualities) (35a and 48b). This latter faculty is the name Descartes gives to that by and through which the mind perceives the wax: "I. . . grasped (the wax) solely with the faculty of judgment, which is in my mind" (33a). Descartes moves on at this point without further explanation of how the faculty of judgment works or more description about what it is and does. So, based on the text, we are forced to conclude that Descartes considers this faculty of the mind absolutely basic. The argument, summed up and simplified, goes something like this:

(1) I know that the ball of wax I am holding now is the same wax I was holding just a moment ago, even though:

(2) All of its qualities discernible by the senses have changed, including its primary qualities.

(3) I know that the wax is capable of more definite changes than I can represent to myself through any imaginative faculty of the mind.

(4) I can only know the ball of wax through *some* faculty of the mind. Let it be called the "intellect."

Since no further insight into the intellect is given, let us move to Descartes's treatment of the ontological question of just what it is that constitutes the ball of wax and is thus perceived by the mind.

Here, Descartes is not much more forthcoming, although he raises the ontological question himself several times in the wax discussion. The closest he comes to answering it is when he says: "Let us focus our attention. . . and see what remains after we have removed everything that does not belong to the wax: only that it is something extended, flexible, and mutable" (32b). However, this is where he begins the argument, noted above, in which the indeterminate number of ways in which the wax's extension and shape can be altered demonstrates the inability of his representational imagination to perceive what the wax is. So it is not clear that even this answers the question. For once he is done critiquing his imagination, Descartes asks yet again, "what is this piece of wax which is perceived only by the mind?," to which he only answers, "it is the same piece of wax I took it to be from the beginning" (33a). But this leaves us with more questions than answers. For example, just how does the mind distinguish between wax and, say, a pen, if it is not by means of sensible attributes? Later, Descartes seems to contradict himself when he says: "I also sensed light, colors, odors, tastes, and sounds, on the basis of whose variety I distinguished the sky, the earth, the seas, and the other bodies, on from the other" (49a-b). Even more problematic, though, is the question of why it just so happens that my mind only ever perceives objects in the world when they are within range of my sense organs? There are many possible questions to be asked. However, I think that part of Descartes's confusion stems from his failure to distinguish and separate out several philosophical questions raised by his discussion. In the next section, I attempt to redress this failure. I think that both the general epistemological (how we know a thing) and metaphysical (what a thing is) questions within Descartes's wax discussion can and ought to be broken down into more specific questions that call for different answers. Although the epistemological and metaphysical questions are related, I will initially break them down separately, considering the epistemological ones first.

Epistemological and Perceptual Questions

Descartes asks how he knows that the ball of wax remains the ball of wax though all of its sensible qualities change. There are multiple questions of epistemology and perception raised here. The first two, which are really two sides of the same coin, I will call the *attachment* or *coherence* problem and the *discrimination* problem. The third is the problem of *change*. I will treat each in turn.

How do I know that certain qualities in my sensory field attach to certain things? According to Descartes, all the sensible qualities of the ball of wax change. To be more precise about the implicit phenomenology here, we might say that at time t1 I was presented with a certain amalgam of sensible qualities and that later, at time t2, I was presented with another, wholly different amalgam of qualities. Why would I judge that these qualities were meant to go together in a particular way? Why should they *attach* or *cohere*[3] into a particular thing, and the *same* thing, no less, each time? If all that the senses give me are qualities or properties, why do I say that they are qualities of a particular thing, entity, object, etc., and how am I justified in doing so? The fact that what we call a "thing" can undergo such changes in its sensible qualities is supposed to motivate the question of why we should attribute any set of sensible qualities to *a* particular thing.

When performing his ball of wax experiment, Descartes was certainly aware of other qualities besides those attributed to the ball of wax. Indeed, he speaks of his hands, a fire, etc. So then, he did not perceive *one* thing but several, and discriminated between them. Not only, apparently, does my throng of sensations attach or cohere to one thing, the sensations are also appropriately differentiated and associated with different things. But, again, why should this be so? Why and *how* do I discriminate between things in my sensory field? If I consider, for example, my visual field, why does it appear to be populated with different *things* and not, rather, just a smear of colors—some changing (or, rather, disappearing and being replaced), and

[3] I use more than one word to avoid any commitment or implication to a stance about the metaphysics of the things perceived.

others remaining static? This is what I call the discrimination problem.

Finally, the ball of wax raises the problem of how we perceive change and, especially, changes in *a* particular thing. Why *do* we say that the ball of wax is still there? How do we perceive changes in our sensory field *as changes in a thing*? These three problems—the problem of coherence/attachment, the problem of discrimination, and the problem of change—are the basic ones I find raised by Descartes's discussion of the ball of wax. However, if we take the perspective of looking at all three questions as a whole, we might further inquire about the nature of the content in our perceptual experience. For example, is our perceptual content primarily conceptual? This is an issue that has provoked a lot of discussion in recent years and which I will consider toward the end of the paper.

If we consider some of the most fundamental things that we do in perception—namely, perceive discrete and particular *things*, discriminate between things, and perceive changes in things—I think Descartes's discussion of the ball of wax shows us something that is indeed correct—namely, the inability of something like a bare experience of pure sensations to accomplish perception. This is why Descartes says that we perform these tasks through a faculty of the mind. But this merely defers the problem and does not solve it. Descartes provides no explanation of how this faculty of the mind performs the task, offers no mechanism through which it might do so, and does not explicitly recognize that it must perform these different tasks. This last point is an important one. Descartes ends his discussion of the ball of wax by only stating that *a* particular faculty perceives the wax. But, as I have been trying to show through my discussion, there are several basic things that we are doing when we perceive the ball of wax. Why should there not be different faculties to perform the different functions? Our bodies seem to have different senses for detecting different qualities of things. In other words, there are different faculties for different things. What reason do we have for believing that only one faculty of the mind, one that Descartes tells us very little about, performs all these different functions? In reality, Descartes's "intellect" is simply a placeholder for a solution to problems he has raised but not sufficiently addressed. But, there are more

questions to be considered—metaphysical ones about the essence of the ball of wax. I will now turn to these.

Metaphysical Questions

Descartes asks just what the ball of wax is after supposedly showing that it remains essentially the same thing through changes in all of its sensible qualities. As quoted above, Descartes answers that the ball of wax is something "extended, flexible, and mutable." Next, I show that Descartes's answer is insufficient and that it is so because the philosopher fails to distinguish between certain, distinct questions about the essence of the ball of wax. I do this by showing that there are at least three ontological senses to the question: what is the ball of wax?

If we make explicit the coherence/attachment problem raised above, we may see that it leads to the general question: just what *is* a thing? If we recognize this, we may ask what the ball of wax is *qua thing*. But calling it extended, flexible, and mutable cannot answer this question. Presumably minds are things as well, at least for Descartes, since they are included in his ontology. But, according to Descartes, minds are not extended. Therefore, "extension" cannot be part of the answer to the question: what is a ball of wax *qua thing*? Now, whether or not Descartes intended to answer this question, it is worth noting that his discussion of the ball of wax does in fact open up the question of what a thing is. While Descartes separates mental and physical things into two different ontological realms (respectively, the *res cogitans* and the *res extensa*), one may still observe that both ontological realms are populated by individual things and thus ask the most general metaphysical question that would cover both realms: what is it to be an individuated thing? However, if we grant that Descartes was not trying to answer this question, we may still ask whether his answer to what the ball of wax is is satisfactory for some other question.

There are at least two more senses to the question of what the ball of wax is. They are: what is the ball of wax *qua res extensa*?, and, what is the ball of wax *qua ball of wax*? The reason we need to distinguish these

two questions is because the ball of wax could both serve as a particular example of a *res extensa* in which we may discover the essence of extended things (namely, their extension), *and* be considered, *in essence*, an extended thing. This establishes the importance of the first question. The importance of the second question is seen in the fact that we must ultimately ask what distinguishes the ball of wax from all other things. To say that it is extended does not answer this question (by this I mean that extension is a necessary but not sufficient condition to the essence of the ball of wax *qua* ball of wax). Unfortunately, Descartes does not adequately answer either of the questions listed above. If "extended, flexible, and mutable" is an answer to the first question, it may be objected that not all extended things are flexible. If it is meant to be an answer to the second question, it may be objected that the list of properties does not distinguish the ball of wax from countless other things in the world. Therefore, though at least three basic metaphysical questions may be asked about the identity of the ball of wax, Descartes's answer is sufficient for none of them. Furthermore, while there may be other important metaphysical questions about the ball of wax, it is hard to imagine Descartes's proposed answer as to "what it is" to be sufficient. So, given that Descartes has failed to answer the metaphysical questions raised by his discussion of the ball of wax, we may also ask if he has answered the epistemological ones. If he cannot tell us what the ball of wax *is*, nor how we can know it (i.e. perform the perceptual functions described above), should we trust any other conclusions he draws from the ball of wax discussion?

Apart from his answers to the particular epistemological and metaphysical questions outlined above, Descartes draws other implications from his meditation on the wax that are important for his philosophy and the course of philosophy ever since. Next, I would like to extend the implications of the critiques I have been developing thus far in order to counter the further philosophical work Descartes thinks his discussion has accomplished for him, as well as add a few more objections to his discussion in general.

Further Objections

First we must consider the possibility that Descartes uses the wax discussion to imply that all our ideas of things are innate. In Meditation Five, in which he says he will investigate the essence of material things, Descartes says he will do this by inspecting the ideas he has of them in his mind. The ideas he considers—like the ones he gave in response to what remains in the ball of wax (i.e. extension, flexibility, mutability)—are primary qualities: extension ("the things quantified in length, depth, and breadth"), "shapes, number, movement, and the like" (45a). Concerning these ideas, Descartes says something deeply reminiscent of Plato's doctrine of recollection:

> Their [(the ideas listed above)] truth is so open and so much in accord with my nature that, when I first discover them, it seems I am not so much learning something new as recalling something I knew before-hand. In other words, it seems as though I am noticing things for the first time that were in fact in me for a long while, although I had not previously directed a mental gaze upon them. (45a)

Based on this account, then, let us consider the possibility that these ideas are innate. Recall that, according to the arguments put forth in the meditation on the wax and recounted here in Meditation Five, these ideas of primary qualities are both beyond the ability of the bodily senses to adequately grasp, *and* they are sufficient for identifying the essence of things like wax. However, if my arguments above are correct, Descartes's proposed answer to the question of the essence of the ball of wax is confused and does not provide sufficient knowledge for identifying the wax. Now consider the requirements Descartes lays down for having innate ideas (or any ideas at all) that we may be sure of: clarity and distinctness. And further consider the amount of clear and distinct knowledge Descartes claims for our other "innate" ideas: God and the soul or mind. Throughout the *Meditations* Descartes articulates all kinds of knowledge he thinks he has about the attributes of God and, concerning the mind, he even begins the second meditation saying that the human mind is known far better than the

body. If we consider both the confusion in Descartes's answers concerning the wax and the relative paucity of knowledge Descartes's claims for them in comparison with other "innate" ideas, I think we are forced to conclude that his ideas about the ball of wax do not pass his own test of clarity and distinctness, and are thus not innate.

A further consequence Descartes seems to draw from his meditations on the wax is the almost complete philosophical insignificance of the bodily senses. Indeed, both his introduction to the discussion and his transitional phrases immediately following it indicate that he thinks the ball of wax discussion puts the final nail in the coffin of the philosophic importance of the senses. And indeed, the senses and the body in general have been relatively neglected in the history of philosophy (at least until recently and even then, only among certain groups). So my next argument aims at reversing this sentence on the bodily senses. I will argue that it *is* (at least conceptually) possible for us to judge that a thing is the same as a previously perceived object *through perception of its sensible qualities*, even if all of those qualities change. I will do this by arguing that the phenomenology of the perception of the ball of wax that I think implicit in Descartes's discussion and that I outlined above is incomplete and unnecessarily excludes a real and important possibility.

The way the perceptual judgment described above could happen through the senses is through the fact that, if I am constantly perceiving the thing through its changes, it would be possible for there always to be at least one identical attribute (though not always the same one) that I perceive with my senses while it changes. All that is required is that there be differential rates of change in the attributes of the thing. Let us assume that as I bring the wax close to the fire, its shape immediately begins to warp, but its color remains the same. Next, the color changes but I notice that the smell is still the same. Note that even when the last *initial* attribute changes (say, its taste) it is still *possible* that *whatever* color (smell, shape, texture, etc., or some combination thereof) it has before that change will remain the same, and I would be justified in judging that it is still the same thing. It is possible that this could happen with two or three attributes. That is, for example,

scent and color could alternate being the constant attribute across successive instants. Or, a combination of attributes, such as color and texture, could remain the same across successive instants. As long as there is always one attribute of the wax that remains the same in two successive moments, it is possible for me to always judge that it is the same piece of wax even though all the attributes eventually change from their initial conditions. Of course, proving that this is a possibility does not prove that this is how it happened for Descartes or how it may happen for anybody. But Descartes never says anything in his account of the wax that could prove it did not happen that way. And the fact that we often seem to perceive things without explicit consciousness that we are doing so (think, for example, of the experience of going on "auto-pilot" while driving and suddenly finding that you have traveled a distance or even maneuvered without being explicitly aware of all or even any thing for which, on reflection, perception appears necessary in order to have accomplished the task) guards against the possible objection of the implausibility of being able to attend to all the ongoing changes versus consistencies from moment to moment. But the important point here is that this "differentially shifting attributes" argument demonstrates that Descartes's meditation on the ball of wax has not definitively proven the default of the senses and imagination in being able to perceive an object. Thus, Descartes is not yet entitled to appeal to some further indefinable simple faculty of the mind that mysteriously perceives the essence of the wax through all changes.

The next argument I want to consider against Descartes's epistemological understanding of the ball of wax might be called the argument from "insufficient skepticism." In the discussion of the wax Descartes never considers the fact that he might simply be mistaken that the same ball of wax remains the same. For someone who entertains doubts about the possibility of having a body and who makes persuasive arguments about the possibility that we might be dreaming or are constantly deceived by an evil genius, it is surprising that he would raise no doubts that the wax is still the same and only offers social convention as his reason: "no one denies it" (32a). However, I think it is perfectly plausible to doubt that the wax remains the

same. In fact it may be a set of historically contingent social practices that lead him to judge that it is the same.

As an illustration of this, consider the bread and wine of the Catholic Communion. Historically, after the priest blessed the bread and wine, it just *was* the flesh and blood of Jesus Christ. That is, mundane objects, like Descartes's ball of wax, made the sort of ontological transformation Descartes seems unwilling to acknowledge as a possibility. Next, consider the case of a bee-keeper who sells wax and honey. For him, what Descartes does to the wax in bringing it close to the fire so irreversibly transforms its marketable properties that it really does just become something different and not really the same ball of wax. My primary purpose in considering these two examples is not to argue that the ball of wax does or does not essentially change, but simply to point out an inconsistency in Descartes's method of doubt and skepticism.

Brief Summary and Transition

Thus far I have summarized and considered Descartes's discussion and arguments concerning the ball of wax. Among various critiques, I have argued that Descartes confuses several problems implicit in his discussion that ought to be separated and explicitly articulated. I have attempted to do this for both epistemological/perceptual and metaphysical issues and have tried to characterize them as fundamental and important philosophical problems. Also, through my critiques of the implications Descartes tries to draw from his discussion—namely, my argument against innate ideas, the argument for differentially shifting attributes, and the argument from insufficient or inconsistent skepticism—I have also tried to suggest some possible alternatives Descartes ignores in addressing the questions raised. In the rest of what follows I would like to consider and recommend some approaches to these problems that are characteristic of the phenomenological tradition of the last century. First, I will briefly introduce answers from two phenomenological philosophers to two of the problems that I articulated above, mostly for the sake of introducing the phenomenological style. Second, and lastly,

I will consider in depth some phenomenological and analytic approaches to the problem of the nature of perceptual content outlined above and weigh in on the debate myself.

Different Approaches

Thinkers in the phenomenological tradition provide different answers to the questions considered in this paper. For instance, consider the perceptual problems of coherence/attachment and discrimination discussed above. I earlier indicated that they were really two sides of the same issue, and I will now attempt to bring together both problems into a broader question about the unity of the *thing* itself. Samuel Todes (2001), in *Body and World*, answers the problem of the perceived unity of the thing by developing a phenomenology of the lived body, in which he argues that the felt unity of the active body gives us access to the unity of physical things. In that discussion, Todes argues against the Kantian idea that a transcendental subject *constitutes* the unity of things—thereby barring us from the thing-in-itself—and instead argues that the felt unity of an active body responding to indeterminate needs in its environment gives us access to the unity of things by *conveying* their significance and intelligibility upon them. In this way, Todes argues for a significance that is immanent in the world itself and not original with the conceptual understanding.

Next, consider the most fundamental ontological question articulated above that asks what a thing is *qua* thing. This is a question that recurs throughout the work of Martin Heidegger. In pursuing what he calls the ontological difference (the fundamental difference that obtains between being itself and any particular being), Heidegger thinks that this question translates into the question of the *being* of the thing—something he argues cannot be analyzed into further, simpler things. He criticizes the philosophical tradition for answering the question of what a thing is in just this way (e.g., classic answers stating that the thing is form and substance, a throng of sensations, an amalgam of properties, etc.). If the latter is the way of answering the question of what a thing is, science has proven more successful than

any armchair metaphysics. However, Heidegger thinks that the ontological question of the *being* of the thing is still a real philosophical question, and answers it by taking seriously the way certain things differentially participate in sustaining and conditioning our worlds and the structures of intelligibility that govern them. Ultimately, Heidegger does not think there is *one* answer to the being of the thing, but different ones for different classes of things (e.g., natural things, tools, works of art, technological things, mundane things, linguistic things [like words and languages], existing things [in Heidegger's special sense of that word, e.g., humans], etc.), and much of his later work aims at giving proper phenomenological descriptions to the being of various kinds of things that does not reduce their being to analysis into further entities.

The Nature of Perceptual Content

Now to the problem of the nature of perceptual content. Recently, in a series of articles that in part draw upon the work of the French phenomenologist Merleau-Ponty, Sean Kelly has argued for the existence of non-conceptual content. I will briefly consider his arguments in relation to Descartes's discussion of the ball of wax and show how a better phenomenology of the ball of wax scenario bolsters Kelly's position.

Phenomenologists take issue with Descartes's characterization of perception. Descartes's approach to the problems outlined above rests on the assumption that in sensation we are primarily confronted with abstract properties. This can be seen in his implicit phenomenology of the perception of the ball of wax that I articulated above. According to this account, at time t1 we are confronted with an amalgam of properties, and at a later time t2 we are confronted with a different amalgam. Thus arises the question of how we perceive the same ball of wax at times t1 and t2. A more recent manifestation of this idea is seen in philosophical talk about "being appeared to greenly," etc. The point is that it is assumed we primarily sense general properties that may attach to any indefinite number of objects. Phenomenologists like Merleau-Ponty and Kelly deny that perception is primarily like

this. Instead, they argue that if we inspect our perceptual experience closely, we rather primarily experience *things* with *aspects* (non-universal qualities of things that are not abstractable from those things themselves) [See Kelly 1999]. The problem becomes one of how we abstract from object-specific aspects to universal properties.

Now, before going on, let me try to dispel a possible misunderstanding in my critique of Descartes and so preempt any criticisms of inconsistency after what follows. I do not mean to criticize any intent in Descartes's analysis of the ball of wax to show that we must bring something to perception. If that is all he meant to show by saying that we must have something in our minds that perceives the ball of wax through its changes, I have no quarrel with him. The problem is that I do not think his answer is that innocent. Descartes's answer seems to do two important things for him. One, he uses his answer to try and dodge the role that the senses and, by extension, the body plays in enabling and constituting our experience of the world. Second, he intellectualizes the perceptual faculty that we have, whatever it is, and makes it overly conceptual. For instance, recall that elsewhere in the *Meditations*, Descartes argues that the same faculty of intellection that perceives the ball of wax is also that which outstrips the imagination in letting us understand that there could be geometric figures with a thousand sides. He argues that we can *conceive* (through the use of concepts like numbers) more than we can *imagine* (through the use of mental images). But, by assimilating the perceptual faculty with that which allows us to conceive of things like unimaginable geometric figures, Descartes ignores the possibility of non-conceptual content that is nonetheless meaningful in our perceptual experience. It is to this possibility that I now turn.

Kelly has argued in several places for the existence of non-conceptual content in experience (see Kelly references in bibliography). Often, his arguments are directed against those of John McDowell, who argues against non-conceptual content primarily because of his epistemological project for which such non-conceptual content would be very damaging. McDowell has argued that in order to avoid skepticism, we must make sense of the world rationally regulating our beliefs. He thinks that this is only possi-

ble if our experience of the world is conceptual throughout (see McDowell 1994). Kelly has argued for non-conceptual content on phenomenological grounds, quite independently of any denial or advocacy of epistemological doctrines—McDowell's or anyone else's. Let me briefly describe two of Kelly's recurring arguments for non-conceptual content.

Kelly endorses and develops the view he finds in Merleau-Ponty that we primarily perceive things with aspects (not universal properties) [See Kelly 1999]. Besides qualities that are specific to objects, he also thinks we perceive qualities that are specific to certain contexts, especially those of lighting contexts. In the latter case he discusses the phenomenon of color constancy—the fact that we perceive colors to be the same throughout changes in lighting. For example, we see the same grassy green throughout the field, though some spots are in the bright sun and some are in the shade of a tree. However, though the color remains the same grassy green, it often has a different "look" to it in different lighting contexts. Because of this, Kelly says: "Without a reference to the context we won't have the resources necessary to explain the change in experience that occurs when the lighting context is varied" (Kelly, 2001: 607). He goes on: "If it is right, as all perceptual psychologists agree, that this change is not a change in color. . . then no color concept, not even a demonstrative one, could completely describe the content of a color experience" (607). As for the former case hinted at above, Kelly also argues that certain qualities are specific to the object they are found in. So, for example, he quotes Merleau-Ponty saying, "the blue of the carpet would not be the same blue were it not a woolly blue" (607). The same would go for the "grassy green" mentioned above. The point, according to Kelly, is that experience is filled with so many context-specific and object-specific aspects and qualities, that it is highly implausible that we experience them all conceptually (even through demonstrative concepts [e.g., "this shade"]—the strongest argumentative tool employed by McDowell against non-conceptual content). Elsewhere, Kelly sums up and defends his argument against the demonstrative concept countermove:

> It is still open to the conceptualist to argue that perceptual content is explicable in terms of the conjunction of a variety of

> demonstrative concepts—one that picks out the property, one that picks out the object that manifests that property, and then a large set of demonstrative concepts that picks out the relevant features of the context in which the property is being perceived. But it seems as though this last set will present a sticking point, since there could be an indefinitely large number of relevant contextual features, and which features of the context are relevant will change from situation to situation. This seems to me the real reason that perceptual content is non-conceptual—because it's situation dependent, and situations aren't specifiable in conceptual terms. (10-11)

Another proponent of non-conceptual content, Christopher Peacocke, in a paper somewhat critical of Kelly's arguments, has presented a different argument for non-conceptual content. His argument is somewhat more of a naturalistic, or evolutionary, one. He says that while we would resist attributing the possession of concepts to certain "lower animals" we would not deny that we share at least some common perceptions with them. As he says: "If the lower animals do not have states with conceptual content, but some of their perceptual states have contents in common with human perceptions, it follows that some perceptual representational content is nonconceptual" (Peacocke, 614). And he ends his article with the following, "Nonconceptual representational content is part of our animal heritage" (615).

Peacocke thinks that Kelly's arguments are not decisive against the conceptualist. For example, while Peacocke says he is sympathetic to treating matters of differential illumination as non-conceptual, he thinks the conceptualist might as yet be undefeated in being able to hold that "... anyone who enjoys the distinctive experiences of color constancy has to have some at least rudimentary concept of illumination" (612). He thinks further work is needed to settle the issue. However, while Peacocke may show that Kelly does not have a knock-down argument against the conceptualist, I think there is a deficiency in his own account. If Peacocke's argument succeeds, it merely proves the fact *that* there is non-conceptual content in perception.

It does nothing to describe or show how that non-conceptual content figures into our experience, how it might interact with the conceptual content, how meaningful it might be, if or how we might respond to it in action, whether it influences our intentional states such as propositional attitudes, and so on. Kelly and Merleau-Ponty's phenomenological approaches, on the other hand, have the virtue of trying to do just that. Certainly it is difficult, perhaps even paradoxical, to try to describe and characterize non-conceptual content in conceptual terms. However, if we were to accept something like Peacocke's argument in proving the existence of non-conceptual content, Kelly and Merleau-Ponty's approaches will take us much further in answering the kinds of questions indicated above. In what follows, I try to bolster Kelly's claims by attempting an alternative phenomenology of the heated ball of wax, one that I think offers insight into the questions raised above.

The problem with Kelly's account is that the philosopher considers cases of perception that are largely static, overly "spectatorial," and do not fully capture the perception of the *active* body using all of its senses. Because of this, he is forced to focus on qualities like colors—hallmark examples of universal properties—that are more susceptible to conceptualist critiques. But, in the case of the heated ball of wax, Descartes has kindly offered us a case of temporally extended perception where moving bodily interaction with a changing thing is the key component under inspection. When I perceive the ball of wax as I move it toward a fire, I experience its changes as *appropriate* to it. The ball of wax solicits from me a certain bodily readiness for the types of changes it will undergo.[4] I perceive and respond to these changes smoothly and fluidly, even though I cannot entirely say or specify conceptually what all those changes are or will be. I can say that I should be utterly shocked if the wax turned a bright pink and grew sharp spikes in an instant, but I cannot put a conceptual limit on what will surprise me and what not, nor can I entirely specify, even with demonstrative concepts, all the different aspects and changes that I attend to, am ready for, and cope with. Furthermore, I perceive that the wax could

[4]For more on body-sets and readiness see the references by Dreyfus, Todes, and Wrathall.

change in slightly different and indeterminate ways than in those with which I am currently presented. I perceive that the ability of the wax to do this outstrips my mental imagination, but, interestingly enough, also my conceptual faculties. Unlike geometric figures where my ability to conceive possibilities has a one-to-one correlation with the number of sides (concepts) I can come up with, I perceive that the wax admits of more changes in its various properties (color, smell, shape, texture, etc.) than I can either imagine or conceive. Hence, here is another argument against Descartes's idea that the primarily conceptual "intellect" is sufficient for our perception.

Finally, consider the possibility, outlined above, of perception of differentially changing attributes. If it so happens that we do indeed identify the wax through perception of identical qualities through successive instances (even though the beginning and ending states do not share identical attributes), it also seems that we are unable to conceptually pick out just what those specific attributes are from moment to moment. If this is indeed the way that the perception of the ball of wax happens, then this represents strong evidence for non-conceptual content.

Above I raised the issues of how non-conceptual content figures into our experience, including our conceptual experience. I indicated that a phenomenological approach, rather than a purely logical one, or one about scientific findings, was best equipped to handle them. I think I can now further cash out that claim. In one of his essays, Kelly expresses sympathy with the epistemological concerns motivating McDowell's conceptualism. He suggests that we should not rule out the possibility of non-conceptual content being able to serve as the basis for thought, though he confesses that he has no idea how such a possibility might play out (Kelly, 2001a: 418-9). However, if my account of the perception of the heated ball of wax is correct, it counts as an instance of something in the world *normatively*[5] influencing our conceptual talk about it through non-conceptual content. Recall that, phenomenologically, we experience the ball of wax as constraining our bodily readiness

[5]Kelly, building on Merleau-Ponty's notion of maximal grip, discusses a different type of normativity associated with the perception of objects (see Kelly 1999). While my particular use is different, I certainly credit Kelly's notion for helping inspire mine.

for it and thus giving us a sense of what kinds of changes are appropriate to it, even if we cannot fully spell out that appropriateness conceptually. However, that sense for appropriate change surely figures into our shared convictions with Descartes that the wax "remains the same." Thus, we have a case of non-conceptual content normatively influencing our conceptual beliefs—and in a way that seems completely rational.

To summarize, my phenomenology of the perception of the ball of wax attempts to uncover non-conceptual content therein. It is argued that such content can be found in (1) the active body's sense of the appropriateness of the changes occurring in the wax, (2) the perception that the wax may continue to change in indeterminate and not fully conceivable ways, and (3) the differentially changing aspects/attributes that perception uses to identify the wax through change (assuming it does so, absent any better account). Furthermore, it was argued that the normative, yet not fully conceivable, influence of the wax on our perception of its ability-to-change must also influence our conceptual talk about it, thus indicating a possible instance of non-conceptual content influencing judgment that we would all accept as rational. If my argument succeeds, it should also recommend the further use of phenomenological approaches to the questions raised in this paper, as they do more than analytic arguments such as Peacocke's to describe and give a place for non-conceptual content in our experience of the world. Finally, another consequence of the success of my argument would be wider consideration in philosophy of cases of active perception in which there is more change, movement, and the coordination of the bodily senses working together in concert, rather than simply passive, or primarily visual and static perception.

Works Cited

Ariew, Roger. and Eric Watkins, eds. *Modern Philosophy: An Anthology of Primary Sources*. Indianapolis: Hackett Publishing Company, Inc., 1998.

Borgmann, Albert. *Technology and the Character of Contemporary Life*. Chicago: The University of Chicago Press, 1984.

Carman, Taylor, and Mark B. N. Hansen, eds. *The Cambridge Companion to Merleau-Ponty*. Cambridge: Cambridge University Press, 2005.

Descartes, René. "Meditations on First Philosophy." trans. D. Cress. *Modern Philosophy: An Anthology of Primary Sources*. Indianapolis: Hackett Publishing Company, Inc., 1998.

Dreyfus, Hubert L. "Introduction I: Todes's Account of Nonconceptual Perceptual Knowledge and Its Relation to Thought." *Body and World*. Cambridge, MA: The MIT Press, 2001.

Heidegger, Martin. *Poetry, Language, Thought*. trans. A. Hofstadter. New York: Harper Colophon Books, 1971.

Kelly, Sean. "The Non-conceptual Content of Perceptual Experience: Situation Dependence and Fineness of Grain." *Philosophy and Phenomenological Research*. 62 (2001): 601-8,

———. "Demonstrative Concepts and Experience." *The Philosophical Review*. 110:3 (2001a): 397-420.

———. *The Relevance of Phenomenology to the Philosophy of Language and Mind*. New York: Garland Publishing, Inc., 2001b.

———. “What Do We See (When We Do)?.” *Philosophical Topics*. 27:2 (1999): 107-128.

———. “What Makes Perceptual Content Non-conceptual?” *Electronic Journal of Analytic Philosophy*. 1998. <http://ejap.louisiana.edu/EJAP/1998/Kelly98.html, 1-12>.

McDowell, John. *Mind and World*. Cambridge, MA: Harvard University Press, 1994.

Merleau-Ponty, Maurice. *The Phenomenology of Perception*. C. Smith, trans. New York: Routledge, 1958.

Peacocke, Christopher. “Phenomenology and Nonconceptual Content.” *Philosophy and Phenomenological Research*. 62 (2001): 609-615.

Todes, Samuel. *Body and World*. Cambridge, MA: The MIT Press, 2001.

Wrathall, Mark. “Motives, Reasons, and Causes.” Carman, Taylor, and Mark B. N. Hansen, eds. *The Cambridge Companion to Merleau-Ponty*. Cambridge: Cambridge University Press, 2005.

Interview With John Perry, Stanford University

John Perry is Henry Waldgrave Stuart Professor of Philosophy at Stanford University and a founder of the Center for the Study of Language and Information. He is preeminent in the philosophies of language and mind. His many important publications include *Situations and Attitudes* (with Jon Barwise), *The Problem of the Essential Indexical*, and *Identity, Personal Identity and the Self*. Perhaps most important, he is famed for his warmth, his genius, and his wit. He co-hosts *Philosophy Talk*, the popular philosophy radio show out of San Francisco, California. The interview was conducted via email over the week of April 11, 2005.

THE YALE PHILOSOPHY REVIEW: *What prompted you to study philosophy?*

JOHN PERRY: When I was a kid I was interested in religion and philosophy. I read Will Durant's *The Story of Philosophy* and found it quite fascinating. I especially liked Schopenhauer, and that led me to an interest in Eastern Religion for a while. It never occurred to me that philosophy could be one's profession until I went to college—Doane College in Crete, Nebraska, a small liberal arts school. I quickly decided that I wanted to be a college professor since it seemed like such a pleasant life. I liked literature, history, and even mathematics, but philosophy was what I liked best.

YPR: *Why have you chosen the philosophical subjects you have studied? Is there an overarching theme or question that you have pursued?*

JP: My choice of topics has been haphazard and opportunistic. By the

time I graduated from college, I had developed an interest in Wittgenstein, so I went to Cornell, where Max Black and Norman Malcolm were leading lights. They had both been students of Wittgenstein, and there was a lot of interest in Wittgenstein at Cornell. Pretty soon I had the *Philosophical Investigations* pretty much memorized, like many of the graduate students at Cornell in the mid-1960's. (It was quite a group, including Bob and Marilyn Adams, George Wilson, Chuck Marks, Ken Walton, and Ron Giere, among others.) But I became more interested in Sydney Shoemaker and Keith Donnellan and the topics they were discussing than studying Wittgenstein. Timothy Smiley, Jonathan Bennett, and H.P. Grice also visited Cornell while I was there. So my interests tended to follow whatever the most exciting seminar was being offered. I ended up more involved with Carnap, Quine, and Frege than with Wittgenstein. I went to UCLA in 1968, and quickly became interested in what David Kaplan, David Lewis, Richard Montague, Tom Hill, Warren Quinn, Montgomery Furth, John Taurek, and others were doing there.

The last few years, however, I have been returning to the most basic philosophical problems—things like consciousness, time, and free will. I suppose I feel that with time slipping away, I'd better try to take my best shot and dealing with these problems.

YPR: *It has been twenty-two years since the founding of the Center for the Study of Language and Information (CSLI). In your opinion, what has been the most significant accomplishment of the center?*

JP: CSLI was formed by a consortium of research groups from Stanford's philosophy, linguistics, and computer science departments and groups in computer science and AI at Xerox's Palo Alto Advanced Research Lab and SRI. The Stanford psychology department became involved soon after the founding.

Jon Barwise, John Etchemendy, Stanley Peters, and I were all working on situation semantics, and it would be pleasant to suppose that the development of our ideas was the most significant thing. Probably, however,

the most important developments coming out of CSLI were in linguistics, where there were two groups, known then as LFG and GPSG, the first involving Stanford's Joan Bresnan and Ron Kaplan (at Xerox), the second Ivan Sag and Tom Wasow. Although various versions of Chomsky's framework are probably still most popular among linguists, the CSLI-based theories have had an enormous impact in computational linguistics, and I don't think Stanford's linguistics department would have risen to its present high stature without CSLI.

Etchemendy and Barwise went on to do a lot of important work on the Liar Paradox, visual reasoning, and have also had a large impact on the teaching of logic. I hope you are using *Language, Proof and Logic* at Yale. Your wonderful professor Sun-Joo Shin was also a product of CSLI's research environment, with its emphasis on visual reasoning and diagrammatic proofs. My favorite CSLI-influenced product is the theory of information David Israel and I developed based on the situation-theoretic concept of constraint, and my own "reflexive-referential" theory, also based on ideas that grew out of situation semantics. I can't say that these things have changed the world, although there is still time and hope. Finally, a number of more practical projects at CSLI have been important; I especially think about the Archimedes Project, which has to do with information and disability.

Along with CSLI we started an interdisciplinary undergraduate major, Symbolic Systems, which Jon Barwise led in its early years, and Tom Wasow has been overseeing most of the time since then. The major has been very successful, with graduates playing important roles in most of the famous companies that have come out of the Silicon Valley in the last twenty years. No doubt the undergraduates, graduate students—like Sun-Joo Shin, Mark Crimmins, Stephen Neale, Paddy Blanchette, and Leora Weitzman—and the post-docs who have been able to pursue interdisciplinary interests are the most important result of CSLI.

Finally, I want to mention the *Stanford Encyclopedia of Philosophy*, which I hope you are all familiar with. Ed Zalta came to Stanford as a postdoc years ago, and has stayed around developing two impressive things:

his own extremely powerful theory of abstract objects and ontology and the *Stanford Encyclopedia*. These are both housed in the Metaphysics Research Lab at CSLI.

YPR: *Philosophy seems to lend itself well to interdisciplinary projects because philosophical concerns seem to ground every discipline. But do experts in other fields have some sort of responsibility to reflect philosophically on their area of expertise?*

JP: I think there are two or three questions here. First, should people in disciplines like linguistics, psychology, computer science, and literature be aware of some of the basic ideas and distinctions that philosophers make? I have in mind particular work by analytical philosophers. I think the answer is clearly yes. We do manage to come up with useful ideas, terminology, and distinctions. Others should learn them.

People in other disciplines should be careful—as we in philosophy should be and often are—in coming to firm convictions about big philosophical issues, however. Based on a weekend's dabbling in biology, I wouldn't expect to have anything very profound to say about genetics. So, second, I would hesitate to think that people in other disciplines had any responsibility to reflect philosophically on their area of expertise, in the sense of thinking that they are going to have much of interest to say to philosophers. Some will, some won't.

Of course, people ought to think about the ethical implications of what they are doing, and avail themselves of the good body of work that has been developing in "practical ethics" to do so.

So, in sum, people in other disciplines should educate themselves about the ideas and distinctions philosophers have to offer; they should certainly tell us when our thinking is over simplistic or out of date of just stupid with respect to their areas of expertise—politely, if possible. They should certainly ponder the ethical implications of the things they do. Those who think that, now that they have mastered some real discipline, it is time for them to take a few minutes and solve a philosophical problem or two, should

mostly go watch TV instead.

YPR: *Tell us about* Philosophy Talk. *Why did you decide to do it? Do you enjoy it?*

JP: It's bothered me for a long time that philosophy seems to play such a small role in public discussion, and when it does it seems mostly perceived as a matter of studying dead philosophers. Among the most prominent philosophical voices are people like my brilliant Stanford colleague Richard Rorty, who has a pretty gloomy vision of analytic philosophy and what it offers. I think philosophy is an enjoyable activity for many intelligent people, that philosophical discussion and reflection, although it solves few problems on its own, has much to offer on almost all topics of importance, and many of little importance for that matter. On public radio one can hear intelligent discussions of history, literature, music, and politics. Why not philosophy?

Finally, I am a fan of *Car Talk*. Now, people aren't really all that interested in cars, but they find humorous animated discussion about automatic transmissions and how odors get into air conditioners and things like that engaging. Philosophy is intrinsically more interesting than automobiles, and most philosophical problems are less complicated than the average automatic transmission. So it seemed to me a program about philosophy, even with less humor and less interesting personalities than Click and Clack, should work.

It's a lot of fun. Ken Taylor is the perfect colleague and co-host. We get to talk to lots of interesting people, too. By the way, *Philosophy Talk* is carried live on KALW radio, 91.7 in San Francisco. You can hear it live over the Internet, <http://www.sfusd.k12.ca.us/programs/kalw/index.html>. And you can hear it on other radio stations at various times, and you can stream past shows over the internet any time you want. We hope to have it available for podcasting soon. For information about all of this, you can start with my website <http://www-csli.stanford.edu/ john/> and follow the links. Or you can go direct to <http://www.philosophytalk.org>. There are

links to KALW and to the *Philosophy Talk* blog, "cogito ergo blog."

YPR: *What have been your first reactions—positive or negative—to the new blog?*

JP: My own blogging has been off-again on-again, depending on how far behind on my teaching I am and how much time my family, which includes three children and ten grandchildren, two cats and four goats, takes up. Ken is a much better organized and more disciplined person, and blogs regularly—some of his posts are really excellent. We are getting some good input from people like Alex Nehamas and John Fischer, who have been on the program. Some good comments from Yale students would be much appreciated.

YPR: *Have you seen any benefits to* Philosophy Talk *beyond your own satisfaction?*

JP: Well, I think it is having some impact. Once we are polished enough to go national it'll have more. Does Yale have a station that would like to carry the program? I know of at least two other philosophy programs, one in Australia and one, "Guerilla Radio," which is student-run, at UC Santa Barbara, which we may have had a role in inspiring. So maybe philosophy's presence on the radio will grow and grow. I hope so.

By the way, let me put in another plug for the *Stanford Encyclopedia of Philosophy* here too: <http://plato.stanford.edu/>. We are trying hard to keep the encyclopedia free and university libraries all over the world, including many third-world libraries, are helping with this effort—although not yet Yale's, last time I looked.

YPR: *How do you answer someone who says, "What's the point of philosophy?"*

JP: I think philosophy is first and foremost a natural human activity and

interest, at least for some people. If you sat in your sixth-grade classroom and wondered whether your fellow students were automatons or real people, you are a natural-born philosopher, and you lead a happier life if you have some chance and encouragement to think about such issues. From the point of human enjoyment of the life of the mind, philosophy has as much point, or as little, as literature, art, and other things humans enjoy.

More practically, if one traces back most of the great changes in human life we count as progress, such as women's liberation, the ideal of equality, and political freedom, one will find at some point the articulation of ideas, ideals, and visions of the possibilities for human life by philosophers played some role. And so too, in fact, for many scientific and technical developments, such as logic and computer science. So it is of some practical value.

Idealists think that ideas constitute or create the world. If so, they are certainly important. But even if this is not so, ideas, concepts, principles, metaphors, dogmas, and doctrines of all sorts shape our experience and practice. Philosophy's job is to examine these things, sometimes to clarify them, sometimes to distinguish them, sometimes to debunk them. It's a noble calling, and also a lot of fun.

YPR: *Why is Personal Identity an important subject?*

JP: The concept of a person, and of the same person, is at the heart of how we think of ourselves and the world around us. It is involved in our ideas of responsibility, freedom, property, virtue, vice, and practically everything else. When we think about the conditions of personal identity, we are led into issues about the nature of thought, consciousness, memory, and other fascinating things, not to mention time itself, perhaps the most slippery of all philosophical problems.

YPR: *Someone says, "X and Y are the same F, but they're* not *the same G." How would you respond?*

JP: Well, it depends on X and Y and F and G, I suppose. Suppose someone says, "That over there (pointing at Berkeley College) and that over there (pointing at Calhoun College) are the same university but not the same college." I'd say they were no doubt correct. If they went on to say, "That shows that there is no such thing as identity, because the two things both are and are not identical," I'd say they were wrong. Berkeley and Calhoun are different colleges, but parts of the same university. If they would put their original insight more carefully, and say, "That college over there, and that college over there, are not identical, but the university of which that college is a part is identical with the university of which that college is a part," then identity wouldn't be threatened.

The interest in examples of "same F, different Gs" is due to the brilliant philosopher Peter Geach. I call the phenomenon he was getting at the "relativity of individuation." Our concepts organize things in lots of different and overlapping ways: the same area is a college and a university, the same lump is a bit of clay and a statue, the same office is occupied successively by different people, and so forth. Of course there was Heraclitus, too, who noted that the same river might not be the same water. He drew the conclusion that you couldn't step in the same river twice, but I think that was a mistake; you can step in the same river twice, but you may be stepping in different water.

Geach called the phenomenon the "relativity of identity" and drew the conclusion that there was no such thing as identity of the sort Frege and Quine believed in—a universal relation that obtains between each thing and itself and no other thing, and obeys the laws associated with "=" in classical logic. This is explained in his book *Reference and Generality* and elsewhere. In my first article, "The Same F," I argued that the relativity of individuation doesn't force us to accept the relativity of identity, since all examples can be handled in some way analogous to the one above about Berkeley and Calhoun. My article was required reading on the topic for a while, but now there is a pretty huge literature. I still like my article.

I think what I said in "The Same F" was right, as far as it goes, but I have been increasingly impressed with the extent to which we get by with-

out fully individuating what we are talking about. Think about clouds. We can talk about clouds just fine, for the most part, without having to develop a clear picture of where one cloud leaves off and another begins. The question of at what point discourse forces more and less explicit individuation is an interesting one, which I say a little bit about in *Identity, Personal Identity and the Self*.

YPR: *What do you think about the idea that personal identity is just what we care about?*

JP: Parfit argues that what we should care about, and to a certain extent—faced with a range of examples of fissioning of fusing persons—do care about, is a relation, psychological continuity, that doesn't meet the axioms of identity. If he's right, it won't do to say that personal identity is just what we care about, because what we care about might not, for example, be transitive. Perhaps Methuselah at 300 cared about what life would be like for Methuselah at 350, and Methuselah at 350 cared about what life would be like for Methuselah at 400, but Methuselah at 300 couldn't have cared less about Methuselah at 400.

I'm inclined to think that even if personal identity is not the same as identity of the live human animal—a conclusion we might be driven to by considering brain or body transplants, brain duplications, Star-Trekian "beaming" and other not-quite-everyday phenomena—still our best bet for understanding why we care about our future selves in the special way we seem to is to think of it first from a biological point of view, and then examine how cultural constructs like our concept of person, responsibility, property, and the like add to this. Why do I care about the fact that my head aches in a different way than I care about the fact that your head aches? Why do I anticipate the root canal I will have tomorrow in a different way than I think about the root canal you will have tomorrow? Why do I save for my old age but not your old age? These are three somewhat different questions. At the bottom is why we do or should (if we should) care about anything at all. The Schopenhauer in me says that it is a vile trick by nature,

the Nietzsche in me says it is nature's greatest gift, the Hume in me says, "Quit thinking about it and get some lunch."

YPR: *Over the years that I've had my car, I've renovated it piece by piece and so completely that not one piece remains from the original car. Do I own the same car? How do you propose to solve this type of problem?*

JP: Well, even if your mechanic saved all of the pieces and reassembled another car built from them, you own the same car and he owns a different car. That's what the law says (I think), and it seems right to me. On the other hand, if the car you did this with was one of historical importance, say the car JFK was riding in when he was shot in Dallas, the mechanic's car will be of more historical interest, and probably of more monetary value, than yours. If we are clear about parts and wholes, properties and dates, the problems will work out. Your car had the property of JFK sitting in it in 1963, and so has the property that JFK did sit in it in 1963. Your mechanic's car doesn't have either of these properties, because it didn't exist in 1963. On the other hand, his car has the seat in it that JFK sat in in 1963, and yours doesn't. That property of his car might be of considerable interest to assassination scholars and JFK buffs. The strange thing that happened to your car means that the observable, examinable properties it has don't tell us as much about the properties it had as we might have otherwise expected, and the strange history of your mechanic's car means it has properties that are of much more historical interest than those of most cars that are only a couple of months old.

My own practice is to keep cars for about twenty years, replace as little as possible, and not allow politicians to ride in them. So I avoid these kinds of problems.

YPR: *Tell us about your recent work in Free Will. Why does it interest you?*

JP: When I was an assistant professor at UCLA, John Taurek convinced me to be a compatibilist: free will and determinism are compatible. I've taught

the problem in introductory classes for almost forty years, occasionally dipping into the literature. It's pretty clear that the arguments for compatibilism don't convince everyone, including some really smart people. Should they really convince me? I'm no longer sure. I'd like to understand the problem better. My present conjecture is something like this. We first need to understand why Fatalism is wrong. Fatalism says that simply the fact that any proposition is true or false is enough to dislodge freedom. If I'm going to eat a pastrami sandwich for lunch, then the proposition that I eat a pastrami sandwich for lunch is true. So the proposition that I would eat a pastrami sandwich for lunch today must have been true yesterday; it had to be either true or false, and it can't have been false. But if it was already true yesterday that I was going to eat a pastrami sandwich for lunch today, how can I do anything today that will change that? Changing the future seems to involve changing the past. But we can't change the past.

Now it seems until I really understand what is wrong with that argument, we can't understand what additional problems the thesis of determinism adds for freedom. And if we don't understand that, we don't understand why getting rid of determinism doesn't solve the problem.

The more I think about it, the more I think that our attitudes towards time are playing an important role. This is a truly frightening thought for me. It has been hard enough for me to learn what I know about classical logic. If I become convinced that it harbors some misconceptions, will I have to learn about three-values logics? Or will I have to learn what relativity theory teaches about time? How about tense-logic? Am I up to learning about these things? It's not at all obvious.

Still, I have tenure, a wonderful job where I am paid to think about such things, and wonderful students and colleagues to help me. So I am plunging ahead on this problem, surely one of the most basic and most difficult of all philosophy.

By the way, this illustrates something important about philosophy. In the end, it really doesn't matter if others have figured this out already. Quite likely Taurek had it all clear. Still, philosophical problems are like mathematical proofs: you have to work it out for yourself to really take care of

the philosophical itch and feel you understand things. Of course, in mathematics, the experts can tell you which proofs are correct, so that you should try to understand them. We seldom have that luxury in philosophy.

YPR: *Questions of Free Will are often raised in theological discussion. Does theology interest you? If so, have your thoughts on theology influenced your thoughts on identity?*

JP: I have always found the questions religion takes seriously interesting, even though it's a long time since I counted as much of a believer. Thinking about immortality, whether it is possible and whether it is something I would want, has played a big role in my philosophical ruminations, even though I haven't thought it is at all likely that it will happen. I am also inclined to think our attitudes towards time have a religious origin. In particular, the "view from no when," the "God's eye view," which I am inclined to think makes no sense, seems to have a role in the way people who are not at all religious think about time. So, yes, theological issues play a role in my thinking, although mainly in a negative way.

As to religion, my current view is that the proper object of human worship is the earth, or the Earth, I guess, along with other aspects of nature; a form of paganism, really. What we need to do with religion is not get rid of it, for its interests and commitments and the human emotions it captures are too important. We need to think about how best to harness them, or, to put it in a less self-important sounding way, to find out what the proper objects of human worshipful attitudes are. I think worshiping an all-perfect being, even if there is one, is a diversion.

Special thanks to the Yale Philosophy Department, the St. Anthony Hall Education Fund, Chitra Sundaram, and the Yale Undergraduate Funding Committee for their support.

This issue of THE YALE PHILOSOPHY REVIEW was produced using MiKTEX <http://www.miktex.org/> and TEXnicCenter <http://www.texniccenter.org>, and was printed by Turley Publications, Inc. of Palmer, Massachusetts. The typeface is Times.