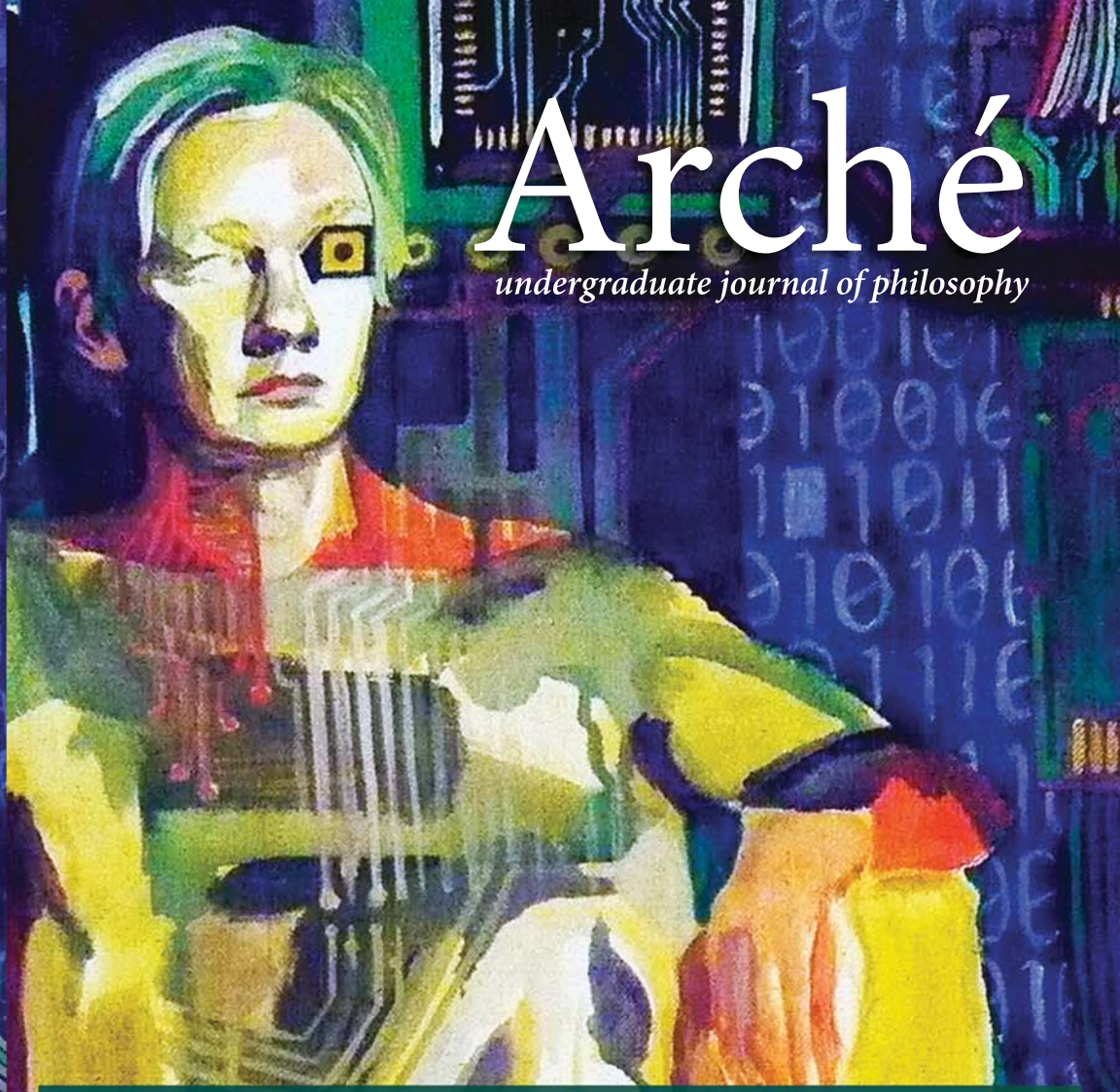
arché

journal of philosophy
volume 5



Arché

undergraduate journal of philosophy

Volume v, Issue 1: Winter 2012
Published at Boston University

In Defense of Stoicism
Alex Henderson

Locke on Nominal vs. Real Essence, and the Identities of Substances:
Why a Mass of Matter and an Oak Tree Can Be in the Same Place at the Same Time
Michelle Mary Dyke

A Defense of Epistemic Responsibility
Corey Cusimano

Reconciling Eros and Neuroscience:
Maintaining Meaningful Expressions of Romantic Love in a Material World
Andrew Pellitieri

ἀρχή

Arché: An Undergraduate Journal of Philosophy
Volume V, Issue 1: Winter 2012

CONTENTS

3 *Letter from the Editors*

4 *In Defense of Stoicism*
ALEX HENDERSON

15 *Locke on Nominal vs. Real Essence, and the Identities of Substances: Why a Mass of Matter and an Oak Tree Can Be in the Same Place at the Same Time*
MICHELLE MARY DYKE

32 *A Defense of Epistemic Responsibility*
COREY CUSIMANO

60 *Reconciling* Eros *and Neuroscience: Maintaining Meaningful Expressions of Romantic Love in a Material World*
ANDREW PELLITIERI

83 *About the Contributors*

ORDERING. Back issues are available for $10.00 per copy. Additional charges may apply for shipping overseas. To place an order and arrange payment, please contact the Managing Editor at buarche@gmail.com.

SUBMISSIONS. Undergraduate authors are invited to submit all types of philosophical discourse: research papers, critical analyses, conference proceedings, interviews, translations, letters and commentary on texts previously published in *Arché*, and original philosophical works. Submissions and correspondence should be sent as Word or .rtf attachment, with "Submission" in the subject heading, to buarche@gmail.com. Submissions may also be sent by mail to: Undergraduate Philosophy Association, c/o Boston University Department of Philosophy, 745 Commonwealth Avenue, Boston, Massachusetts 02215.

Please include a 100-word abstract. The name of the author(s) should appear on the first page only. Citations should follow Chicago style, with all references occurring as endnotes. Authors are encouraged to query if they would like assurance that their submission is of a kind usually considered by the editors, or to follow-up if their submission has not been acknowledged in a timely manner.

ISSN: 1946-1801
Cover art: "Twenty-first Century Schizoid Man" by Marguerite Mooradian.
Printed by Offset Prep, Inc., of North Quincy, MA.

Letter from the Editors

Productive philosophy requires critical thinking and self-examination—capacities that would not be possible without active engagement of other perspectives in discourse. This is a familiar lesson taught to us early on by Socrates, who thought that people made judgments far too hastily—that they referred back to traditions and authorities before consulting their own minds and engaging others. Since its inception, philosophy has been built upon the element of discourse; without an ongoing dialogue among thinkers of different creeds and cultures, philosophy would be fruitless and stagnant, unable to push new boundaries or arrive at new conclusions.

In this issue, we have four papers which exemplify the fruits of philosophical discourse. Alex Henderson offers us new insights to Cicero's intentions, while Michelle Dyke challenges a Locke scholar's interpretation. Philosophy further expounds that we need to have sense beyond our own perspectives. Corey Cusimano questions the extent to which we are epistemically blameworthy and responsible for our own beliefs. And lastly, Andrew Pellitieri challenges the views of the Churchlands, searching for the presence of romantic love beyond a material understanding.

This issue of *Arché* carries special significance, as it is a product of the work of last year's editorial staff in cooperation with this year's—a dialogue among the old and the new, with seasoned perspectives and fresh eyes coming together to produce an outstanding issue. We hope that the message of this journal, and of the papers herein, will remind its readers that no man is an island, and that we must come together and share our opinions in order to do something truly great.

Kelly Kearney
B.A. Philosophy, 2012

Jenna Kreyche
B.A Philosophy, 2011

with special thanks to Richard Cipolla

In Defense of Stoicism

by ALEX HENDERSON*
Georgetown University

This paper takes Cicero's assault on Stoic philosophy from *Pro Murena* as a point of departure to engage three critical aspects of Stoicism: indifference to worldly concerns, the sage as an ideal, and Stoic epistemology. It argues that Cicero's analysis fails to clearly distinguish between these elements of Stoic philosophy and, therefore, presents Stoicism in a misleadingly unfavorable light.

Charges of bribery were brought against Lucius Licinius Murena in 62 BCE. Despite entrenched opposition from the popular party, Murena was able to enlist the aid of the famed orator and presiding consul, Marcus Tullius Cicero. His prosecutor was the most conservative of senators—Marcus Porcius Cato, famed for his rectitude and his unbending adherence to Stoic philosophy. In order to defend Murena, who was almost certainly guilty, Cicero chose to go on the offensive and discredit his opponent by undermining Stoic philosophy in the eyes of the jury. Cicero portrays Stoicism as follows:

> A wise person never allows himself to be influenced… Philosophers are people who, however ugly, remain handsome; even if they are very poor, they are rich; even if they are slaves, they are kings. All sins are equal, so that every misdemeanor is a serious crime… The philosopher has no need to offer conjectures, never regrets what he has done, is never mistaken, never changes his mind.[1]

Cicero's portrayal seems to argue that (1) the Stoics' commitment to remain indifferent to worldly influence causes them to lack compassion for the circumstances of other people, (2) Stoics are too severe when confronted with the indiscretions of others because the Stoic theory of justice does not distinguish between severe crimes and mi-

* arh52@georgetown.edu. Received 1/2011, revised December 2011.

nor infractions, and (3) Stoics tend to be absolutist in their convictions and refuse to hear the other side of a situation once they have made up their mind on an issue. By describing Stoicism in this way Cicero was able to portray Cato as an overzealous advocate of the letter of the law, implying that the prosecution brought the case to the court in order to align with philosophical dogma rather than out of concern for the common good. The neglect of Hellenistic philosophy in the modern academy has combined with Cicero's dominating literary personality to allow this attack to go largely unanswered by Stoicism's defenders.[2] This paper will argue that Cicero's speech distorts Stoicism for persuasive purposes and obscures the true tenets of that philosophy. Cicero's critique will be used as a point of departure for discussing three fundamental Stoic doctrines: indifference to worldly concerns, the regulative ideal of the Stoic sage, and Stoic epistemology.

The Origins and Rationale of Stoic Indifference

Cicero argues that the Stoics believe "the wise person never allows himself to be influenced… only fools and triflers show mercy. A real man is not affected by prayers or attempts to placate him."[3] Here he implies that Stoic autonomy is so rigorous that it dehumanizes its practitioners and prevents them from having compassion for their fellow men. In the context of the trial, he contends that the prosecution is incapable of exercising the "kindness and humanity" that should prevail in the judgment of Murena.[4] The Stoic doctrine that a man should be indifferent to all but his own vice and virtue traces its origins to Socrates and admits of greater complexity than Cicero allows. It will shed light on the matter to trace the development of this doctrine and work out some of the implications that follow from it.

Stoics have a clear position when it comes to externals (those things which are not in the direct control of an individual's moral purpose or character); they believe them to be 'matters of indifference' (*adiaphora*), that is, neither good nor evil.[5] This position is best summarized by Diogenes Laretius:

> The Stoics say that some things that exist are good, some are bad, and

> some are neither good nor bad. The good things, then, include the virtues… The bad things include their opposites… neither good nor bad are all those things which neither benefit nor harm.[6]

The Stoics trace this doctrine back to Socrates. In the *Apology*, Socrates famously asserts that:

> Neither Meletus nor Anytus can harm me in any way; he could not harm me, for I do not think it is permitted that a better man be harmed by a worse; certainly he may kill me, or perhaps banish or disenfranchise me, which he may and maybe others think to be a great harm, but I do not think so.[7]

The indifference of Socrates to the attacks of the wicked is explained in the extended argument of the *Gorgias* and culminates in Socrates' final indictment of Callicles:

> All the other theories put forward in our long conversation have been refuted and this conclusion alone stands firm, that one should avoid doing wrong with more care than being wronged, and that the supreme object of a man's efforts, in public and in private life, must be the reality rather than the appearance of goodness… Be guided by me then and join me in the pursuit of what, as our argument shows, will secure your happiness both here and hereafter. Let people despise you for a fool and insult you if they will; nay, even if they inflict the last indignity of a blow, take it cheerfully; if you are really a good man devoted to the practice of virtue they can do you no harm.[8]

If Socrates' arguments are taken seriously, then bodily harm, poverty, shame, and anything else that is not directly related to a man's character cannot hinder him in fulfilling his moral purpose and living a happy life. The Cynics were the first of the Socratic schools to take this aspect of the Socratic paradigm to heart, and it is through them that the Stoics trace their relationship to Socrates.

Antisthenes, the founder of the Cynic school, studied under Socrates (as did Plato) but emphasized very different aspects of So-

cratic teaching. Cynics focused chiefly on Socratic ethics rather than epistemology and metaphysics. This led them to adopt rigorous asceticism and applied ethics as the keys to living excellently.[9] Epictetus, Stoicism's most influential proponent in the Roman Empire, viewed the Cynics as a sort of living testimony to virtuous happiness, proof that "god has sent us to show that it is indeed possible" to live a happy life bereft of worldly prosperity.[10]

Though they admire them, the Stoics argue that Cynic asceticism is too extreme. In shunning wealth, honor, etc., Cynics place a negative value on them. This contrasts with the teaching of Socrates, who in the *Gorgias* described such ends as "intermediate and neutral," as things which acquire moral value through their proper or improper use.[11] For Stoics, externals should neither be shunned nor sought, but accepted as they come. Epictetus gets at this point in the *Enchiridion* when he says, "For this is your duty, to act well the part that is given to you; but to select the part belongs to another."[12] Cynics, by cutting themselves off from civilized life, refused to accept the roles required of them in human society. To the Stoic way of thinking, the Cynics are like actors in a drama who take the play so seriously that they refuse to play the part assigned to them because they scorn costumes and scripted dialog as ostentatious. They fail to realize that these things have no real value one way or another except insofar as they are used in the production. By obsessing over props and scenery, the Stoics argue, Cynics miss the larger production and fail to play their roles fully. In this light, Stoicism's doctrine of indifference to externals is seen to be a restoration of the initial Socratic position on the relationship of the philosopher to worldly concerns.

The leading figures of Stoicism were, as Stanford Professor A. A. Long aptly puts it, "interpreters and embodiments of the Socratic paradigm."[13] Instead of focusing on the intellectualism of Socrates, the Stoics, like the Cynics, are chiefly interested in Socrates as "a new kind of hero, a living embodiment of philosophical power, a figure whose appeal to the Hellenistic world consisted in self-mastery."[14] The Stoics conceived of indifference to outside influences as a prerequisite to achieving self-mastery and the freedom of choice necessary to live a virtuous life. The Stoic writings are filled with injunctions to this ef-

fect, from the extensive discussion of slavery, freedom, and externals in the first part of the *Enchiridion*, to Seneca's reminder that "Liberty is not to be had gratis; if she be worth much to us, all things else have little value."[15] The Stoics allow for the possibility of total freedom from evil by denying, along with Socrates, that evil can happen to them against their will. If good and evil rest entirely within choices between virtue and vice, it follows that the manifold things not subject to personal agency are not, from a first person viewpoint, evil or good. Furthermore, if virtue is the only good and can be had at any time by choosing rightly, then man is free at any time to have the good if he can master the atavisms of his nature that tempt him into error. While Cicero rightly notes that Stoics strive to move beyond the influence of base emotions, he fails to mention that their reason for doing so is to ensure that they treat other people in a just and humane way that is separate from the passions of the moment.

How Stoics Reconcile Indifference with Worldliness

Cicero's critique seems to be that the Stoic is unwilling to take part in the necessary grit of day-to-day life; he uses an improper metric of value and cannot see that the good life requires man to value things beyond virtue. The Stoic reply to Cicero rests on the position that indifference to externals and involvement in the world are not mutually exclusive. Chrysippus makes the decidedly non-anchoritic argument that "they are mad, who make no account of riches, health, freedom from pain, and integrity of the body, nor take any care to attain them."[16] The commitment to indifference is in obvious tension with the Stoic contention that it is reasonable to take account of day-to-day concerns. Cicero's assault rests on this tension snapping the thread of Stoic logic into incoherence.

To reconcile this paradox, the Stoics rely on the doctrine of preferred indifference. Preferred indifference grants that one should strive to be genuinely indifferent to all external things, but, ceteris paribus, human beings naturally prefer some indifferent things to others and are justified in pursuing them so long as they don't allow them to open the door to desire and vice. The Stoics want to maintain that

this preference does not constitute desire in the common sense of the term, but a certain acceptance of the norms and roles that every human being is obliged to abide by in the larger community of mankind. This balancing act is described by Epictetus and revisited by Vice-Admiral James B. Stockdale, an American war-hero who led the prison underground in the Vietcong's most infamous torture camp and went on to write extensively on Stoicism's enduring relevance to the human condition:

> Everybody does have to play the game of life. You can't just walk around saying, "I don't give a damn about health, or wealth, or whether I'm sent to prison or not." Epictetus says everybody should play the game of life—that the best play it with "skill, form, speed, and grace." But like most games, you play it with a ball. Your team devotes all its energies to getting the ball across the line. But after the game, what do you do with the ball? Nobody much cares. It's not worth anything. The competition, the game, was the thing. The ball was "used" to make the game possible, but it in itself is not of any value.[17]

This view is commensurate with the Stoic critique of Cynicism: if externals really are indifferent then it should not matter whether one has wealth, runs a household, or rules a nation—it should matter only that one acts virtuously within the context at hand. It would be strange to deny ourselves worldly goods or sever ties of friendship acquired in the pursuit of virtue if they lack the power to affect us one way or another. Moreover, Stoics hold that it is in accord with nature that man prefers certain basic comforts and relationships, as Seneca says, "Our duty is to follow nature, but it is contrary to nature to torture one's body, to dislike cleanliness and seek filth, or to make use of rough and dirty food."[18] So long as worldly bounty and social engagement do not cause the philosopher to fall into the cycle of desire that leads to vice, the Stoics accept these as concomitants of the human condition.

Preferred indifference allows a Stoic to engage in normal human life—a prerequisite for displaying most virtues—but to remain, in theory at least, indifferent to the actual results of this engagement. They may have wealth, but they should not hesitate to part with it if

honor demands it. Friendships should be cultivated, but companionate loyalty should not be allowed to make us complicit in another's vice. Through these means the Stoics argue that freedom and virtuous happiness need not be bought at the price of a human life. Cicero is incorrect in implying that Stoicism prescribes empathy; rather, Stoicism stipulates that while it is natural to empathize with another person's condition, those feelings should not be allowed to become so strong that they cause the philosopher to abide and abet rank criminality—for example, electoral bribery.

The Regulative Ideal of the Stoic Sage

The last two points of Cicero's critique rest in exploiting the concept of the sage in order to make Stoicism seem both excessively harsh in its treatment of moral failings and stubbornly dogmatic: "All sins are equal, so that every misdemeanor is a serious crime… The philosopher has no need to offer conjectures, never regrets what he has done, is never mistaken, never changes his mind."[19] This characterization conflates a description of a regulative ideal with actual practice.

The Stoic framework of human nature is acutely aware of the vast gulf between what mortal men are generally capable of achieving and what they might achieve in a more perfect world. Tad Brennan, a modern interpreter of Stoicism, explains their position well when he says that, according to the Stoics,

> all of us are awash in vice—everyone that was alive in the time of the Stoics, including the Stoics themselves; every historical figure they knew of, including their most revered predecessors such as Socrates and Diogenes the Cynic.[20]

The Stoics see vice and virtue as fundamental conditions of the soul: on one level individual actions can display virtue or vice, but from a higher vantage point a person is either virtuous or he is not. The truly virtuous man is called the sage, and he is virtuous in all that he does without qualification.[21] The vast majority of human beings are not sages. Therefore, on a fundamental level, they are vicious. This isn't to say

that right action is frivolous, on the contrary:

> The Stoics did allow that virtue was possible for human beings—it is not an unrealizable ideal, just a very demanding one—and they described what it would be like to make progress towards virtue. They even allowed that some people do make progress. They simply denied that making progress towards virtue was the same thing as becoming more virtuous or less vicious.[22]

Cicero is right to claim that, in light of the unity of all vices, the Stoics believe that "the man who unnecessarily strangles a cock is as guilty as the man who has strangled his father,"[23] but only in the limited sense that each action "stems from a vicious state of soul."[24] The man who errs slightly and the man who errs greatly are united in the fact that they have yet to attain the state of real and enduring virtue demarcated by sagacity. Stoics would not say that the crime of killing one's father deserves the same punishment as that of being a negligent farmer, but they would agree with Socrates that "if a man goes wrong in any way he must be punished, and the next best thing to being good is to become good by submitting to punishment and paying the penalty for one's faults."[25] Cicero's argument cleverly but abusively combines the Stoic belief in the universal viciousness of the non-sage with the belief that each man deserves correction for every wrong choice in order to make it seem that they believe all men deserve the *same* correction for every wrong choice. In view of the Stoic framework, the basis of Cicero's accusation can be recognized as a straw man created by blending two distinct positions into a single contention with which no reasonable person, Stoics or otherwise, would have agreed.

Cicero's three contentions concerning epistemology, that the philosopher "has no need to offer conjectures… is never mistaken, never changes his mind," employ a similar tactic. In Stoic epistemology all data that comes by way of the senses, and even the imagination, are impressions. It is within the power of every human being to assent to the impressions they receive, refrain from assenting, or deny the impressions altogether and hold the reverse position.[26] The basic unit of Stoic understanding is true belief, which is a weak assent to an impres-

sion.[27] Weakly assenting to an impression means that the non-sage has accepted an impression under certain conditions, but would alter his assent under some other condition (e.g. under torture or if the assent were shown to be irrational—anything that would cause him to recant). For all non-sages this is the horizon of possible understanding. The sage is a special case in that he has the ability to know something as well as believe it. The sage does this by strongly assenting to a *kataleptic* impression. A kataleptic impression is an impression "with a sort of guarantee… if you are having this impression then things are really as the impression says they are."[28] When a sage strongly assents to a kataleptic impression he recognizes truth and holds fast to it in a way that "cannot be reversed or overturned by any amount of rational questioning, no matter how skillfully conducted, or by any amount of emotional or psychological pressure."[29] Quite importantly, the reason for this tenacity is that the sage is correct in his knowledge of the truth.[30]

The Stoic sage would not recant or be mistaken in his knowledge because Stoic epistemology defines knowledge in such a way as to make either circumstance impossible. Moreover, the sage will not offer conjecture about something that he is not certain he knows because he abides by the Socratic dictum that "Every form of pandering, whether to oneself or to others, whether to large groups or to small, is to be shunned; oratory is to be employed only in the service of right," and will not risk misleading others by persuading them that wrong is right, even out of ignorance.[31] A non-sage practitioner of Stoic philosophy is not the same as the sage. He will mistakenly give assent to false impressions and reverse his position when he is shown his error. His every word will be a conjecture because no matter how true his belief is he will never really *know* anything. As Seneca says: "It is folly to say, 'What I have said must remain fixed.' There is no disgrace in having our opinions change with the circumstances."[32] To pervert this complex epistemology into the claim 'Stoics never admit they are wrong' is disingenuous and does not accurately reflect the Stoic system.

Conclusion

By examining Stoic contentions out of their original context, Ci-

cero successfully undermined Cato's credibility, and won his case—though victory in this case does not mean his arguments were sound. Viewed in light of the actual doctrines of the Stoics, Cicero seems to have won a classic oratorical victory wherein "an ignorant person is more convincing than the expert before an equally ignorant audience."[33] One would be hard-pressed to claim that Cicero was ignorant of Stoicism, but it is fairly clear that the circumstances of the trial prompted him to pander to the jury and play on their ignorance and prejudice in order to facilitate a guilty man's acquittal. The verdict of Plutarch on the character of Cato provides the final vindication of the man and his philosophy and, through example, shows what argument can only intimate: "though he were terrible and severe as to matters of justice, in the senate, and at the bar, yet after the thing was over his manner to all men was perfectly friendly and humane."[34] ■

Notes

1 Cicero, *On Government*, (tr. Grant, New York: Penguin, 1993): p. 143.
2 Martha Nussbaum, "'The Arbitrariness of Canons: The Neglect of Hellenistic Philosophy and Why It Is a Bad Thing.'" <http://www.college.columbia.edu/core/lectures/fall2005>.
3 Cicero, op. cit., p. 143.
4 Ibid, p. 111.
5 Tad Brennan, *The Stoic Life: Emotions, Duties, and Fate*, (Oxford: Clarendon, 2006): p. 119-122.
6 Ibid. 119.
7 Plato, *Five Dialogues: Euthyphro, Apology, Crito, Meno, Phaedo,* (tr. Grube, Indianapolis, IN: Hackett, 2002): 30c-d.
8 Plato, *Gorgias*, (tr. Hamilton, Baltimore: Penguin, 1960): p. 527.
9 Julie Piering, "Antisthenes." *The Internet Encyclopedia of Philosophy*. Web. 02 May 2010.
10 Epictetus, *The Discourses of Epictetus: the Handbook, Fragments*, (tr. Hard, London: Everyman, 1995): p. 194.
11 Plato, op. cit., p. 467.
12 Epictetus, *Enchiridion*, (tr. Long, Amherst, NY: Prometheus, 1991): p. 21.
13 A. A. Long, *From Epicurus to Epictetus: Studies in Hellenistic and Roman*

Philosophy, (Oxford: Clarendon, 2006): p. 9.

14 Ibid, 7.

15 Frederic M. Holland, *The Reign of the Stoics*, (New York: Charles P. Somerby, 1879): p. 89.

16 Holland, op. cit., p. 127.

17 James B. Stockdale, "Stockdale on Stoicism II: Master of My Fate," (Center for the Study of Professional Military Ethics. Web. 31 Mar. 2010): p. 10.

18 Holland, op. cit., p. 126.

19 Cicero, op. cit., 143

20 Brennan, op. cit., p. 36.

21 Brennan, op. cit., p. 36-37.

22 Ibid, p. 36.

23 Cicero, op. cit., p. 143.

24 Brennan, *The Stoic Life*, p. 36.

25 Plato, *Gorgias*, p. 527

26 Brennan, op. cit., p. 60.

27 Ibid.

28 Ibid, p. 68.

29 Ibid, p. 69.

30 It is fairly clear that the notion of a kataleptic impression was created in order to illustrate a larger theoretical point and does not correspond to a particular category of impression that is readily translatable into the parlance of common experience. Like obscenity, one knows it when one sees it, but is otherwise hard pressed to describe it in detail.

31 Plato, *Gorgias*, p. 527

32 Holland, op. cit., p. 133.

33 Plato, *Gorgias*, p. 459

34 Plutarch, Internet Classics Archive, trans. John Dryden, "Cato the Younger."

Locke on Nominal vs. Real Essence, and the Identities of Substances: *Why a Mass of Matter and an Oak Tree Can Be in the Same Place at the Same Time*

by MICHELLE MARY DYKE*
Yale University

Dan Kaufman argues that there is a major tension between Locke's account of identity and his account of the real and nominal essences of substances. I argue that Kaufman's criticism follows from a failure to recognize that properties associated with a nominal essence need not be limited to the instantaneously observable features rooted in one of what Kaufman terms an "individual real essence." Locke's two theories are indeed consistent.

In his 2007 paper, "Locke on Individuation and the Corpuscular Basis of Kinds," Dan Kaufman argues there is a major inconsistency in the views presented in Locke's *Essay Concerning Human Understanding*.[1] Kaufman claims there is a contradiction between Locke's views on identity and his account of the distinction between the real and nominal essences of substances. The real essence of a substance, according to Locke, is the makeup of its particulate parts in nature; this fundamental constitution is not precisely known to us, although it underlies all of the substance's observable qualities.[2] Locke emphasizes that human beings divide substances into kinds on the basis of a variety of observable features. The nominal essence of a substance refers to those properties that are essential to it insofar as it is considered a member of that named kind.[3] To illustrate the alleged problem with Locke's account, Kaufman presents the example of an oak tree and the particulate mass of matter that composes it. From Locke's opening remarks on the identities of "things" in Chapter XXVII of Book II, it

* michelle.dyke@yale.edu. Received 1/2011, revised December 2011.

follows that an oak tree and the mass of matter that constitutes it at a given moment are two different things, and they can coincide in the same place and time provided they are of different "kinds." Intuitively, the different kinds they exemplify are simply those indicated by the names "oak tree" and "mass of matter," respectively. Yet Kaufman believes that Locke's discussion of real and nominal essence, in light of Locke's corpuscular understanding of the basis of qualities, commits Locke to the claim that it is impossible for any *x* and *y* to share the same real essence and yet differ with respect to any of their nominal essences. In other words, no two coinciding substances, like an oak tree and its constituent mass of matter, can be of different kinds.

Yet Locke neither is, nor should be, committed to such a claim. After explaining Kaufman's criticism in more detail in Section I, I argue in Section II that Kaufman's argument fails to take into account that the properties associated with a given nominal essence need not be instantaneously observable features rooted entirely in any one of what Kaufman refers to as an "individual real essence." It is for this reason that two things like an oak tree and the mass of matter that constitutes it at a particular instant can share an individual real essence in that instant and yet be characterized by two different nominal essences associated with two different sets of properties. The alleged tension can be resolved. As I explain in Section III, this realization helps to clarify the entirety of Locke's discussion of the identities of substances. Locke's theories of the essences and identities of substances are not only compatible, but in fact intimately related. This does not mean, however, that Locke's account is wholly compelling. I address some particularly pressing difficulties in Section IV.

I. Kaufman's Challenge

Locke's statements in Chapter XXVII of Book II do indeed commit him to the claim that an oak tree and the mass of matter constituting it at a given moment are two distinct things of different kinds. Locke states that it is impossible for "two things of the same kind [to] exist in the same place at the same time."[4] He also writes that "one thing cannot have two beginnings" in place and time.[5] As Vere Chappell had

already argued in "Locke on the Ontology of Matter, Living Things, and Persons," it follows from these principles and from Locke's own descriptions of the persistence conditions for trees and masses that they are not identical things.[6] A mass of matter, for Locke, is a particular collection of conjoined particles; the same mass cannot survive a loss, replacement, or addition of parts.[7] Yet the persistence conditions for organisms involve continuation of the same life, according to Locke.[8] A tree can survive the loss of a few particles and remain the same tree.[9] It would, however, be constituted by a new mass of matter, since its constituent particles would not be the same. Given his view that no one thing can have two beginnings in place and time, Locke must hold that a tree and the mass of matter constituting it at a given time are not the same identical thing. They must be two things of different *kinds* because they do coincide.[10]

Kaufman refers to Chappell's interpretation of Locke, according to which a mass and an organism are two non-identical coinciding things of different kinds, as the "Coincidence Interpretation."[11] Kaufman writes that in order to defend Locke's metaphysical system as consistent, the defender of this interpretation must be able to account for what the two different kinds are, exemplified by a mass and an organism, that allow them to coincide. The intuitive answer, he agrees, would be that those two kinds are simply mass and organism, respectively.[12] Yet the great difficulty, according to Kaufman, is that Locke's discussion of real and nominal essence does not allow for an organism and the mass constituting it at a given time to be of different kinds at all, because these substances share the same "individual real essence," and no two things of the same individual real essence can differ with respect to any of their nominal essences.

Locke presents his view of essence as an alternative to the classic view held by the Aristotelians, who "suppose a certain number of those Essences, according to which, all natural things are made, and wherein they do exactly every one of them partake, and so become of this or that *Species*."[13] Locke insists instead:

> The other, and more rational Opinion, is of those, who look on all natural Things to have a real, but unknown Constitution of their

> insensible Parts, from which flow those sensible Qualities, which serve us to distinguish them one from another, according as we have Occasion to rank them into sorts, under common Denominations.[14]

There is thus a key distinction for Locke between real and nominal essence. Real essence refers to that unknown constitution of corpuscular parts in nature, from which observable qualities spring.[15] Locke explains that we group the objects of our experience under certain names that emphasize different observable characteristics we take to be important or noteworthy. For example, anything we consider to be a sample of the metal "gold" will display a certain color, malleability, solubility in *aqua regia*, et cetera.[16] We, as human language users, designate these characteristics as the necessary "properties" associated with the nominal essence of "gold." Locke believes there is some corpuscular feature or features at the level of the real essence of a sample of gold that ground its observable characteristics. Yet it is only what we observe of that sample, and not its real essence, that can provide us with a basis for classification.

With this much, Kaufman agrees. However, he interprets Locke's distinction between real and nominal essence to imply that the formation of any nominal essence involves merely picking and choosing from among the observable characteristics already possessed by a substance in light of its corpuscular real essence. Each particular example of a substance in nature, Kaufman explains, has an individual real essence, and all of the properties of any nominal essence that substance could be said to possess must be observable characteristics that derive from that individual real essence's particulate constitution:

> Nominal essences are the "Workmanship of the Understanding" (3.3.14), according to Locke. Among the reasons for this label is that despite the fact that the qualities included in the nominal essence are produced by the individual real essence of a body, it is *we* who decide which qualities to include in the nominal essence. Thus, Locke believes that in the formation of nominal essences, both we and the individual real essence play the crucial roles.[17]

While we have the choice of which qualities to include as necessary properties of a substance's nominal essence, these properties will, according to Kaufman, be limited to a selection of those observable features that are rooted in the individual real essence of the substance. Kaufman explains, "Given Locke's commitment to corpuscularianism, nothing other than the individual real essence of an individual *could* serve as the basis for the observable qualities which go into the creation of the nominal essence of a kind."[18] Given that kinds are a class of nominal essence, Kaufman concludes:

> Necessarily, for any x and y, if x and y have the same individual real essences, then for any kind K, either (i) both x and y are members of K or (ii) neither x nor y are members of K.[19]

This is true, Kaufman argues, because whatever observable qualities would be required for membership in a given kind (as necessary properties of the kind's nominal essence), two individuals *x* and *y* of the same individual real essence will either both display or both fail to display these qualities.

Kaufman finds this result to contradict the claim that an oak tree and the mass of matter constituting it at a given time can be two distinct things of different kinds. Kaufman emphasizes that the oak tree and mass are composed of numerically-identical material parts arranged in a numerically-identical manner.[20] In other words, they share the same individual real essence.[21] Kaufman concludes that the tree and mass should not then be able to differ with respect to any of their nominal essences, although that is precisely what would be required in order to say that they are two coinciding things of different kinds. Kaufman takes this to be a fundamental and "intractable" problem.[22] He concludes there is a deep tension between Locke's views on the essences and identities of substances.

II. A Response to Kaufman

Locke's theories can be successfully defended against this attack. Locke's way of thinking does allow for the possibility that two differ-

ent nominal essences can characterize two distinct, coinciding things. Kaufman's error lies in supposing that the formation of a nominal essence involves merely picking and choosing from among the observable features rooted in some one individual real essence. Locke makes no such claim explicitly, and his description of our "complex ideas" of substances encourages an outright rejection of this supposition. The nominal essence of an oak tree, for example, can include properties such as those describing its growth and behavior over time that are not instantaneously observable and cannot be explained in terms of the constitution of just one individual real essence.

Strictly speaking, phrases like "oak tree" and "mass of matter" directly signify not mind-independent material objects, according to Locke, but ideas: "Words in their primary or immediate Signification, stand for nothing, but the *Ideas* in the Mind of him that uses them."[23] "Oak trees" and "masses of matter" are examples of Lockean complex ideas, and more specifically, substances.[24] Locke explains that the necessary properties associated with the nominal essences of substances are themselves whatever ideas we see fit to associate for the sake of identifying substances as we find useful or desirable:

> The measure and boundary of each Sort, or *Species*, whereby it is constituted that particular Sort, and distinguished from others, is that we call its *Essence*, which *is* nothing but that *abstract* Idea *to which the Name is annexed*: So that every thing contained in that *Idea*, is essential to that Sort. This... I call it by a peculiar name, the *nominal Essence*, to distinguish it from that real Constitution of Substances, upon which depends this *nominal Essence*, and all the Properties of that Sort...[25]

The complex ideas of substances are, according to Locke, built up of simple ideas.[26] These simple ideas include not only primary and secondary sensible qualities such as extension and color, but also a wide variety of active and passive powers.[27] Locke offers as examples the power of gold to be melted, the power of the sun to blanch wax, and the power of loadstone to draw iron.[28] Locke explicitly insists that such simple ideas of powers "make a principal Ingredient in our complex

Ideas of Substances."[29] Yet a power is a feature that may not result entirely from the constitution of any one individual real essence because a power refers to an ability to "make, or... receive any change."[30] Locke explains that when we attribute powers to substances, "the *Power* we consider is in reference to the change of perceivable *Ideas.*"[31] A Lockean idea of power is a simple idea resulting from both sensation and reflection.[32] To develop the idea of gold's power to be melted, for example, would require one to make a series of observations of a changing sample of gold and then to reflect on those observations.

Kaufman takes one individual real essence to be one particular arrangement of corpuscular particles. The exercise of a power might actually require a change from one such individual to another. For example, the idea of an oak tree could include powers like the ability to grow in height, the ability to take up water and nutrients from the soil, and the ability to drop its leaves in the fall. An oak tree remains the same tree after absorbing iron and magnesium from its soil overnight. Yet strictly speaking, the corpuscular real essence of the tree before and after is quite different. Exercise of this power requires a change from one individual real essence to another, and so this power of which we have an idea is not an observable characteristic rooted in one individual real essence. This power can, however, be one of the properties included in the nominal essence of an oak tree, so long as the idea of this power is one of those simple ideas contributing to the complex idea of that substance.

Locke was clearly not averse to understanding simple ideas, not only of powers specifically but even of qualities, to be indicative of change over time. For example, Locke's short list of primary qualities includes not only figure and extension, which can intelligibly characterize objects at given instants, but also motion.[33] A simple idea of motion, developed from experience as Locke's empiricism requires, would not be formed on the basis of an observation occurring at one instant; a Lockean idea of motion must be an idea of something occurring over a span of time. Since even simple ideas can refer to processes occurring over time, the complex ideas of substances built up from those simple ideas may also refer to processes occurring over spans of time. A particular "constant regular motion" is in fact one of

the ideas Locke himself lists as a component of our idea of the sun, which he offers as a prime example of a substance.[34]

Given that an individual real essence is one precise, unchanging arrangement of particles, an individual real essence can be the real essence of a substance only at some specified instant. An oak tree and the mass of matter constituting it at time *t* share an individual real essence only at some one time, *t*. It is problematic for Kaufman's interpretation that an instantaneous individual real essence would not even be capable of displaying one of the primary qualities, namely that of motion. Yet if the properties of substances are restricted to observable features rooted in some one individual real essence, and no one individual real essence can itself possess the quality of motion, then Locke should be unable to describe our idea of the sun as having the quality of a particular motion as one of its constituent ideas, and thus as one of the properties of its nominal essence. If Kaufman is correct, then the *Essay* is even more deeply conflicted than Kaufman claims.

One might attempt to defend Kaufman's interpretation by appealing to the distinction between determinate and determinable qualities; one individual real essence might not be moving in a particular way, but perhaps it has whatever corpuscular features are required for the ability to move. Yet this does not explain how the complex idea of the sun can include a determinate "constant regular motion, at a certain distance from us" as one of its constituent simple ideas.[35] What is important to take away from this discussion of the quality of motion is that Locke makes room for even simple ideas to refer to processes occurring over spans of time; a simple idea of a quality or power need not be descriptive of some sensible object in a given instant. When Locke states that the complex ideas of substances are collections of ideas "observed to exist united together," Locke is not saying that all of the characteristics of a substance must be features existing together in one place and time.[36] The essential properties of a named substance need not all be instantaneously observable features rooted in some one individual real essence of an example of that substance. According to Locke, we carve up the natural world as we observe it into named substances as we please, attributing to these substances any properties of which we can form a simple idea.

The interpretation of Locke defended in this section is consistent with a corpuscular understanding of the basis of observable characteristics. The ideas of many powers will be derived from series of observations made over intervals of time. Each instantaneous observation will be an observation of the primary and secondary qualities that are rooted in the particulate constitution of matter. For example, the observation that a tree has grown a foot in height will consist of a collection of instantaneous observations of the color and figure of the tree. If the color and figure observed at multiple distinct instants are characterized by the proper difference, then the tree is understood to have grown. These instantaneously observable qualities have their basis in the corpuscular features of some individual real essence. Yet a tree one foot and two feet in height are not the same corpuscular individual. To identify a growing tree as one substance and to attribute to it the power of growth is to identify a property of the "tree" that is not rooted entirely in any one of its individual real essences, even though all of the observable qualities enabling us to form the idea of such a power over time must derive from corpuscular features at the level of real essence.

Thus, Dan Kaufman is incorrect when he states that Locke is committed to the claim that, given any *x* and *y* which share the same individual real essence (at any instant), and given any kind K, *x* and *y* are either both members of K or both not members of K. An "oak tree" and the "mass constituting it at time *t*" share, at time *t*, the same individual real essence. Yet the non-identical sets of properties associated with each of their nominal essences need not be chosen from among the qualities observed to result from the constitution of that single individual real essence at time *t*. Locke thinks we form our ideas of substances from whatever simple ideas we wish, even incorporating ideas of change over time into our ideas of substances. All of those constituent simple ideas then become essential properties of the substance's nominal essence; "That therefore, and *that alone is* considered as *essential, which makes a part of the complex* Idea *the name of a Sort stands for*, without which, no particular Thing can be reckoned of that Sort, nor be intituled to that name."[37] The nominal essence of an oak tree includes, for example, powers to perform behaviors over intervals

of time, while the nominal essence of the mass constituting the tree at time *t* need not include any such powers. Note that it is unimportant that these two things do indeed exist for different spans of time; the ideas of different properties belonging to each of their nominal essences are sufficient to individuate them. Because the two different nominal essences of these things constitute their different kinds, we can coherently say, contrary to Kaufman's accusation, that an oak tree and the mass constituting it at a given instant are two distinct things of different kinds, coinciding in space and time. The alleged contradiction between Locke's views on identity and his views on real and nominal essence can be resolved.

III. Locke on Identity, Revisited

The realization that Locke thinks nominal essences do include properties that are not instantaneously observable features rooted in any one individual real essence helps to clarify not only Locke's remarks on spatiotemporal coincidence and kinds, but the entire Chapter XXVII of Book II, on "Identity and Diversity." Locke's conceptions of the nominal essences of substances and of the persistence conditions for individuals over time are tightly linked. Locke realizes that the identity of a substance like an oak tree or a horse cannot consist in corpuscular sameness over time.[38] He explains that a tree is considered the same tree even as it grows from a small seedling to a much larger plant, and a horse remains the same horse throughout its lifetime even while it gains and loses weight.[39] The identity of these substances remains the same, "though, in both these Cases, there may be a manifest change of the parts."[40] Locke realizes there must therefore be some difference between the oak tree, the identity of which remains stable over time, and the mass of matter that constitutes it at some given moment:

> [That difference] seems to me to be in this; that the one is only the Cohesion of Particles of Matter any how united, the other such a disposition of them as constitutes the parts of an Oak; and such an Organization of those parts, as is fit to receive, and distribute nourishment, so as to continue, and frame the Wood, Bark, and

> Leaves, *etc.* of an Oak, in which consists the vegetable Life.[41]

Locke is not arguing here that what makes something an oak tree and not a mere mass of matter is just a precise arrangement of parts that enables life at that moment, as opposed to some other jumbled arrangement of particles that does not, because at any instant, as he notes, there will be one arrangement of mere matter that constitutes the tree.[42]

Locke is instead drawing attention to the differences between our ideas of masses and of oak trees. A mass, according to our idea of it, is simply a configuration of particles at some instant. Yet our idea of a tree includes certain abilities, which Locke calls collectively the "disposition" to "continue" to support the same life over time. The exact parts of a tree may change, but something that remains a tree must have the continued ability to "receive, and distribute nourishment." Continued possession of various powers through time is, as part of the very idea of an oak tree, intimately tied to its identity. As we already saw, Locke takes every constituent idea of our complex idea of a named substance to be a necessary property of its nominal essence.[43] Locke also emphasizes that "Existence it self" is the key to a thing's identity; a thing is the same so long as it continues to exist.[44] We can easily answer the question of what it takes for a named substance to continue to exist by referring to the properties of its nominal essence. That thing must continue to display the qualities and powers associated with the nominal essence applied to it in order to retain its identity as an example of that substance.

Thus for Locke, identity is relative to some nominal essence. Locke considers this to be the key insight of his theory of identity; "such as is the *Idea* belonging to that Name, such must be the *Identity*: Which if it had been a little more carefully attended to, would possibly have prevented a great deal of that Confusion, which often occurs about this Matter."[45] If we accept Kaufman's interpretation of Locke, then we cannot make sense of even the broadest aspects of Locke's theory of identity. According to Kaufman, Locke must hold that the properties of nominal essences are qualities observable at some one instant.[46] If this is true, then Locke should have no reason to say that an oak tree

and the mass constituting it at a given time are different things with different names at all, because their observable qualities in that one instant will be exactly the same. The things identified by the names "tree" and "mass" do indeed have different identities according to Locke, because they have different persistence conditions. Given that Lockean identity is relative to nominal essence (we can only refer to the continued identity of named things insofar as they continue to be of that sort), their identity is easily accounted for when we realize that different persistence conditions are entailed by their different nominal essences, which, in the case of the tree, includes powers governing characteristic behaviors through time, and for the mass, a lack thereof.

It is indeed intuitive to think that a named thing's properties, its persistence conditions, and its identity over time are all intimately linked. When we distinguish between an oak tree and the mass of matter constituting it at time *t*, we do take ourselves to be making a meaningful distinction. We take these things to be different precisely because our ideas of them involve very different conceptions of their behavior over time. We realize that an oak tree will be composed of very different arrangements of matter throughout its lifetime. Yet when we refer to a mass of matter at time *t*, we are referring to something that is unique to one instant of time and which would not remain the same in light of any modification of its parts. Locke's willingness to account for oak trees and the masses constituting them at given instants as different things, thereby entailing they have different nominal essences, different persistence conditions, and different identities, should be considered a merit and not a fault of Locke's account.

IV. Difficulties with Locke's View

While Locke's accounts of the essences and identities of substances do allow for one consistent theory, that theory has at least one major problem. Locke explicitly claims that every component simple idea of the complex idea of a substance becomes a necessary property of that substance's nominal essence.[47] Thus, a substance can no longer be called the same substance as soon as it ceases to display any of that substance's necessary properties. Yet consider, for example, the idea of

a swan. According to Locke's own example, the idea of this substance includes, among other characteristics, "white Colour, long Neck, red Beak, black Legs" as well as "a power of swimming in the Water, and making a certain kind of Noise."[48] Yet it is quite easy to imagine an object that ceased to display any one or even multiple of these characteristics, which we would nonetheless continue to identify as a swan. Imagine a swan that loses its feet in an accident, or has its beak taped shut so that it can no longer make noise. What is unclear is how many of these properties an object must cease to display before it ceases to be a swan. What is truly essential to a swan's continued identity is not so clear and simple as Locke's account would suggest.

One might attempt to salvage Locke's theory by objecting that we have not correctly identified anyone's idea of a swan in general; these are characteristics of some or even most swans, but they are not the essential characteristics that really make something a "swan" in anyone's mind. Yet it is unclear that there actually is even one characteristic that every single example of what we would want to consider a swan has in common, especially if we are looking for a set of characteristics not fully shared by any examples of some other named thing we would want to distinguish as different, like a duck or a flamingo.

It is possible that Locke might concede we do not have any such sophisticated conception of what being a "swan" necessarily requires or entails in all instances; he might, however, deny that we need any such account. One major goal of Locke's explanation of the distinction between real and nominal essence seems to be to show that the names we commonly use to designate various things do not necessarily pick out any perfectly rational, natural kind at all. According to Locke, we identify and name various entities in a somewhat arbitrary way as we find useful, and so the question of how many properties an object must cease to possess before it ceases to be a "swan" need only illustrate that our intuitions about commonly discussed kinds break down in unfamiliar and bizarre circumstances. Yet our strong intuition that a swan with only one leg, for example, is indeed a "swan" provides us with an important reason to reject Locke's claim that all of the properties associated with a nominal essence are indeed necessary properties of any thing to which that name is applied. Locke's theory of the

connection between a substance's identity and its nominal essence is consistent, but in order for the theory to be viable, it would need to be modified with a more sophisticated account of our understanding of the connection between an object's membership in a named kind and its possession of the properties associated with the nominal essence of that kind.

V. Concluding Remarks

While Kaufman accuses Locke of a gross inconsistency with regard to his theories of the essences and identities of substances, his criticism hinges crucially upon a failure to recognize an important aspect of Locke's thinking on these subjects. While Locke did believe that observable qualities are rooted in corpuscular features at the level of real essence, Locke did not hold that the properties associated with a nominal essence must be restricted to the types of instantaneously observable characteristics that can be rooted entirely in the particulate constitution of just one "individual real essence."

This awareness resolves the alleged problem with Chappell's "Coincidence Interpretation" of Locke's remarks on spatiotemporal coincidence and kinds. According to Locke's view, there can be two distinct coinciding things of different kinds so long as they are characterized by two different nominal essences that include non-identical properties. These properties can include powers to perform behaviors that occur over spans of time and require changes from one particulate individual to another. This clarification of Locke's account of real and nominal essence also sheds new light on basic aspects of Locke's theory of identity. For Locke, identity must be considered relative to some nominal essence; one must specify a named individual before one can ask, with respect to the necessary properties of that nominal essence, when that named individual begins and ends its existence. Locke can intelligibly distinguish between two things that share the exact same observable qualities at a given moment, like an oak tree and the mass of matter that instantaneously constitutes it, precisely because the nominal essences of these two things can, in light of their different properties, entail different persistence conditions. Locke's ac-

counts of the essences and identities of substances are not only consistent; they reinforce each other in important ways, even if the emerging conception is not totally compelling.

This clarification of Locke's views raises an interesting and difficult question concerning Locke's ontology. As has hopefully been made clear, Locke is more than willing to say that an individual real essence, as Kaufman understands it, can be associated with different sets of powers. Kaufman assumes that on Locke's view, there can be nothing more to an object, on the most fundamental level of the real essence of the thing, than the instantaneous arrangement of its corpuscular parts. Yet it seems open to Locke to hold that these different powers might be grounded in some kind of fundamental disposition that only manifests itself in the long run. Alternatively, Locke might think there really is nothing more to matter itself than its particulate constitution as it can be described at each instant, and it is only our ideas of substances as they are derived from observations over time that involve the attribution of powers for change to collections of matter. This is a difficult question to resolve. Yet Kaufman is perhaps too hasty to assume that Locke's ontology is one in which nothing but particular arrangements of particles, as they can be described at each instant, may exist at the level of real essence. ■

Notes

1 Dan Kaufman, "Locke on Individuation and the Corpuscularian Basis of Kinds," *Philosophy and Phenomenological Research* 75.3 (2007): 499-534.

2 John Locke, *An Essay concerning Human Understanding*, ed. Peter H. Nidditch. (Oxford, Oxford University Press, 1975), III.vi.6.

3 Ibid., III.vi.2-4.

4 Ibid., II.xxvii.1

5 Ibid.

6 Vere Chappell, "Locke on the Ontology of Matter, Living Things, and Persons," *Philosophical Studies* 60.1 (1990): 19-32.

7 Ibid., 24.

8 Ibid.

9 Ibid.

10 Ibid., 22.
11 Kaufman, 505.
12 Ibid.
13 Locke, III.iii.17.
14 Ibid.
15 Locke is clearly partial to corpuscular explanations of qualities at the level of real essence, although Locke's philosophical commitment to corpuscularianism as the *only* possible characterization of real essence has been challenged, for example by Lisa Downing in "Locke's Ontology," *The Cambridge Companion to Locke's* Essay Concerning Human Understanding, ed. Lew Newman. (Cambridge: Cambridge University Press, 2007), 352-380. Kaufman assumes for the sake of his essay that Locke is committed to a corpuscular understanding of real essence, and this can be assumed for the sake of this paper's argument as well.
16 Locke, III.vi.6.
17 Kaufman, 518.
18 Ibid.
19 Ibid., 527.
20 Ibid., 521.
21 Ibid., 521-2.
22 Ibid., 522.
23 Locke, III.ii.2.
24 Kaufman agrees that these two things should be considered substances, although he notes that certain other scholars have disagreed. Locke constantly refers to other organisms like men, horses, and swans as prime examples of substances (II. xxiii.3, II.xxiii.14).
25 Locke, III.vi.2
26 Ibid., II.xii.1.
27 Ibid., II.xxiii.7.
28 Ibid., II.xxi.1, II.xxiii.3.
29 Ibid., II.xxi.3.
30 Ibid., II.xxi.2.
31 Ibid., II.xxi.1.
32 Ibid., II.vii.1.
33 Ibid., II.viii.9.
34 Ibid., II.xxiii.6.
35 Ibid.

36 Ibid.
37 Ibid., III.vi.5.
38 Ibid., II.xxvii.3.
39 Ibid.
40 Ibid.
41 Ibid., II.xxvii.4.
42 Ibid.
43 Ibid., III.vi.21.
44 Ibid., II.xxvii.3.
45 Ibid., II.xxvii.7.
46 Kaufman, 530.
47 Locke, III.vi.2.
48 Ibid., II.xxiii.14.

Defending Epistemic Responsibility

by COREY CUSIMANO*
Brown University

The claim that we have epistemic responsibilities, and that we can be blameworthy for failing to meet those responsibilities, has recently been challenged by a number of theorists on the grounds of non-volunteerism: the claim that we do not have control over our beliefs. I argue that previous responses to this challenge fail to adequately address the problem and, furthermore, provide ancillary reasons to desire a proper defense of epistemic responsibility. I then argue that, despite earlier challenges, we can be responsible and blameworthy for our beliefs, and that this responsibility is properly considered to be epistemic in nature.

My concern in this paper is to argue for the claim that we are epistemically responsible for at least some of our beliefs. As a consequence of this, I will also be arguing that we are appropriately blameworthy for at least some of our beliefs, and that such blame is properly *epistemic* blame much in the way that, for certain actions, I am *morally* blameworthy. Such a view has most recently been argued against by William Alston (2005).[1] Since then, various epistemologists have attempted to defend some of our deontic concepts (like the use of "oughts") by sidestepping the important challenge of doxastic non-volunteerism. I will argue that these theories either ignore or fail to accommodate the concept of responsibility in their deontic framework.

There are at least two reasons to desire a defense of epistemic responsibility and epistemic blameworthiness. The first, which has been acknowledged even by philosophers arguing against epistemic responsibility, is that *talk* of epistemic responsibilities, duties, and blameworthiness is indeed a widespread phenomenon. This is true not only of our locutions concerning the word "ought" such as "you ought not

* corey_cusimano@brown.edu. Received 1/2011, revised December 2011.

believe such a thing!" or "you shouldn't jump to conclusions!" but also of our *attitudes* and *expectations* of each other as believers. Consider the two following scenarios (which I will continue to use throughout the essay):

> (1) I go to the grocery store and form the false belief that *my partner does not mind if I buy us non-organic meat for dinner*. I have no idea that she had just read a new article about the health hazards associated with purchasing non-organic meat and had thus changed her mind from her usual, ambivalent preferences.
> (2) I speak with my partner right before leaving for the grocery store and she informs me of her preferences. In the meat aisle, I nevertheless form the belief that *my partner does not mind if I buy us non-organic meat for dinner.*

When I return home with meat in (1), my partner may be upset that I purchased meat that she doesn't want and so, among other things it was an unfortunate waste of money. However, in scenario (2) she will likely be upset *at me*.[2] When I explain to her that I thought that she didn't mind which meat I bought she will tell me that I did something wrong.

These kinds of judgments and expectations permeate our everyday lives. Furthermore, we seem to have special kinds of words and concepts for people who are notoriously irresponsible or bad believers versus when their poor beliefs are 'not their fault'. An innocent believer may be *dull-witted*, *unaware*, *uninformed*, *just plain dumb*, or *unintelligent*. However, we also freely recognize those who are *ignorant*, *inconsiderate*, and *neglectful*. These different concepts, and their varied synonyms, rely on and are informed by the idea that we can be responsibly and blameworthy believers.

The second motivation relates to the everyday proliferation of deontic judgments noted above, but is especially concerned with judging each other as moral agents. I take it to be a very good excuse, when someone has done something wrong, to reply that they didn't *believe* it was wrong and therefore should not be punished/rebuked/or so on. Of course, often times they will retort that one *should have* believed that

doing *x* was wrong. However, we only appropriately blame people and punish or chastise others when they are responsible for the negative outcomes of their behavior. If we are never responsible (or cannot be responsible) then our claims that people should have believed certain things, especially in the moral domain, lose much of their potency.

Some philosophers, including Alston,[3] argue that there is a sense in which we can be responsible for some of our beliefs, but that it is not *epistemic* responsibility. Is there a reason to think that, given the concern above, we should prefer to conceptualize our responsibility as *epistemic*? Traditionally, philosophers associate the goal of truth as being the defining feature of the realm of the epistemic. Furthermore, it seems that when a person is forming beliefs about what is right or wrong, their primary goal is discovering and believing the truth (as opposed to fulfilling some other set of goals or duties that could also apply to belief formation). Our duties concerning belief formation may be derivative of broader moral duties or responsibilities, but they are nevertheless epistemic.

The Challenge of Doxastic Non-Volunteerism

In the face of these two motivations, the claim that we are responsible for our beliefs suffers from an apparent critical problem: we do not seem to have any voluntary control over what we believe. Generic responsibility, as it may apply morally, prudentially, or epistemically, seems to require that the agent in question have some control over the outcome or their actions in any scenario before they can be held responsible or blameworthy for what happens. This challenge is related to, but importantly different from, the long-standing principle of ought-implies-can.

For the range of deontological concepts we employ, we can distinguish between those in which something good or bad happened, and those for which we are responsible or blameworthy for. Concerning the former, we may claim something like "he should have gone to an all-boys prep school." This seems true only if there is a way in which they actually *can* (or *cannot*). That is, it makes sense to make these kinds of claims because the boy could have easily been enrolled in a

different school by his parents or that the salt actually does dissolve. Traditionally, it does not seem true, or make sense to say things like the boy "ought to have been a Martian" or "ought to have been born to different parents." This is true of moral "oughts" as well. It seems true to say that "she should be nice to her new classmate" but not of any action that the girl could not possibly perform.

For a subset of these deontic ideals or claims, the agent has effective control over whether the ideal realized. Only for these do we claim that a person was responsible for that outcome (or, in the case of a bad outcome, at least, blameworthy). For instance, we certainly would not *blame* or hold the boy responsible for not going to a different school—he had no control over which school he went to. However, the girl had control over how she acted toward her new classmate and, therefore, can be blamed for being mean. The first kind of oughts can be true and applicable even if they do not have voluntary control over whether we fulfill them, as long as they are something we *can* fulfill, whereas judgments of responsibility and blame necessitate voluntariness.

Given this background, we can understand Alston's challenge (of Doxastic Non-Volunteerism) as follows:

> (1) Generic responsibility or blameworthiness for an action or state requires an agent have voluntary control over that action or state.
> (2) People do not have voluntary control over their judgments (actions) or beliefs (states).
> (C) Therefore, people are not responsible for blameworthy for their judgments or beliefs.[4]

The first premise follows from our discussion of "ought-implies-can" and voluntary action. This premise will be challenged later by several philosophers but let us grant it for now. Premise (2) is the main part of Alston's attack and the one that concerns us now. What we will see is that Alston, and a number of other philosophers (like Jonathan Bennett, 1990), claim that we do not have voluntary control over our beliefs because they are not properly responsive to our will.

Alston invites us to introspect about our experiences of perform-

ing certain kinds of voluntary actions, and to contrast them to our experiences of forming and holding beliefs. Alston notes that we have varying degrees of voluntary control. We have the kinds of actions that we can enact and complete just as soon as we can *will* ourselves to do them; turning one's head or voluntarily blinking are good examples of these kinds of actions. Then there are more complicated actions, like opening a door or uttering a sentence, where it far more likely that we will somehow be interrupted and prevented from completing our action. It even seems like we have voluntary control over certain long-term goals and activities. For instance, I show a kind of long-range control over how clean my house is. I can sweep the floors, do the dishes and laundry, and so on, but it likely that I will have to stop in between these activities. I may get tired or need to eat or take a nap, but it is possible for me to return to my activity and eventually complete it.

Forming beliefs, or failing to form *a certain* belief, does not seem be under our control the way these simple activities are. Unlike blinking or closing a door, we cannot simply will to believe something. Similarly, judging a certain proposition to be true is not something we can accomplish simply by going about fulfilling certain sub-goals (like cleaning the dishes). In these examples the only necessary prerequisites to accomplishing something were willing it and a relatively stable environment (i.e. one that would not interfere with the activities set in motion by our will). Judgments of the truth or propositions, or states of belief and disbelief, are formed. Take, for instance, what I consider to be an absurd proposition:

> (a) My television has the power to elongate its cords and strangle people while they are asleep.

It is obvious to me that I could never get myself to believe this proposition.[5] It is also obvious to anyone who attempts to form that belief that it is impossible to do. There are many other kinds of attitudes I may take toward this proposition. I may hold (a) as a hypothesis or factor it into my decision making or even hope, wish, or pretend that it is true, but the attitude of *belief* is impossible to willfully create.

Not only do we seem unable to willfully to form specific beliefs, but the beliefs that we do form are also outside our control. Alston writes:

> When I see a car coming down the street, I am not capable of believing or disbelieving this at will . . . when I look out my window and see rain falling, water dripping off the leaves of trees . . . I form the belief that rain is falling willy-nilly. There is no way I can inhibit this belief.[6]

From this introspective exercise we conclude that beliefs are neither the kinds of things we can willfully induce in ourselves nor the kinds of things we can inhibit when they occur.

Alston concludes that if we have any kind of control over our beliefs, it is *indirect*. Although my beliefs are not responsive to my will or agency, I can come up with a long term goal involving selective exposure to certain kinds of evidence, and so to achieve a specific belief. Some long-term projects, if they were reliable, would entail some kind of effective control over our beliefs. However, the chances for success with such a plan, according to Alston, are likely very slim. Take the proposition (a), above. I may try to watch horror films about televisions, draw angry eyes above the screen, and tell my friends to move our TV to my bedside while I sleep, so that when I wake up, I am (hopefully) startled. But even after all of this effort, my success is not guaranteed. Although I can attempt to influence my beliefs in those ways, I am not guaranteed to achieve my goal and, in fact, most likely will not succeed as a result of my willful actions. That is, I do not have an *effective choice* as to whether I believe something or not.[7] Effective choice of the outcome of a judgment is, for Alston, what it takes to exhibit voluntary control and, therefore, to be responsible or blameworthy.

Some philosophers have appealed to a dichotomy between practical and epistemic reasons as an explanation for why we lack voluntary control over our beliefs. For instance, Jonathan Bennett argues that our will, the source of voluntary action or the idea of actions/states originating from the agent, is the capacity to be responsive to *practi-*

cal reasons.[8] Alternatively, beliefs are responsive to epistemic reasons, where epistemic reasons are those based solely on the truth or falsity of a proposition. The reason why the proposition that *my television can strangle people* is so strange is because there is so little evidence for it. These two kinds of reasons correspond to two different capacities, willing and believing, and normally they do not interact.

There are different ways in which our practical reasons, desires, or interests do influence our beliefs. For instance, if I am highly interested in baseball, I will form beliefs about the game, various teams and players, and so on. In this sense, my beliefs are sensitive or responsive to my interests. The claim made by Bennett and Alston, however, is that even if we can influence generally what we form beliefs *about*, we still cannot control what we believe about those things. That is, we can influence the range of propositions we form judgments about, but not whether each individual proposition is true (or false).[9] So, despite the fact that I can choose to form a series of beliefs (of undetermined content) about the inside of a room (by deciding to walk through it), I still have not exhibited the relevant voluntary control over my beliefs because I cannot decide to believe a specific proposition. There are plenty of times when the involuntariness of beliefs does limit what our beliefs are about (i.e. beliefs about rain or whether a car is speeding toward you), but the relevant limitation is on the *specific* content of the belief (whether it *is* raining or *not* raining, and so on).

Finally, there is one other way in which the dichotomy between "will" and "belief" seems to break down. Wishful thinking, and other forms of self-serving or delusional beliefs, are apparent counterexamples to the claim that our practical reasons do not influence our beliefs. Consider the case of a person who is in love with someone else and desperately desires that person to love them back. Upon being rebuked, that person, instead of forming the belief that that the feeling of love was not reciprocal, might come to believe that their love really was returned, and that the apparent rebuke was nothing other than a test. Such scenarios seem, *prima facie*, to be cases of our desires, or practical interests, influencing what we believe.

However the processes of rationalization and evidential appraisal operate, they nevertheless seem to be unresponsive to our will (even if

they are things we do want). Let us suppose that our best explanation for the situation above is such that our agent treats the other's behavior as evidence *of his or her love* much the same way we treat seeing a car as evidence of a car to us. In this way, we can see how the belief is still involuntary. Although our practical interests or desires may influence our beliefs in this indirect way, we can still see a division between the will and our practical reasons, and beliefs and evidential reasons. Our voluntary actions are still constrained by practical reasons and we are unable to will any indirect or direct influence on our beliefs (or how we appraise evidence for certain propositions) with any sort of effectiveness. From these limitations, and the premise that we must have effective choice/voluntary control over actions or states in order to be responsible for them, we maintain our conclusion that we cannot be responsible for our beliefs or judgments.

Various theorists, when faced with the challenge of doxastic non-volunteerism, have responded in one of two ways. They have claimed either that deontological judgments do not require voluntary control to be true, or they deny premise two, and claim that we actually do have voluntary control over our beliefs. The problem with the first kind of response is that, even if when it succeeds, it jettisons concepts of *blame* and *responsibility*, or at least declaws them. The problem with the second response is that it fails to address the actual challenged issued by doxastic non-volunteerism above.

Rejecting Ought-Implies-Can

Richard Feldman (2004),[10] Hilary Kornblith (2001),[11] and Matthew Chrisman (2008)[12] provide accounts of epistemic deontologism on the basis of rejecting the first premise of Alston's challenge as explicated above. They each neatly build on each other's theories and I will present them in turn. Alston assumes that deontic terms like "should" and "ought" are necessarily intimately tied to their effectiveness as injunctions and our responsibility for responding to them. Feldman, Kornblith, and Chrisman (FKC) have responded by arguing that we can still use plenty of our deontic terms without relying on the premise of voluntary control over belief.

To motivate rejecting the ought-implies-can principle, Feldman shows us that, for plenty of deontic locutions, we already do so. For instance, we frequently say things like "Teachers ought to explain things clearly. Parents ought to take care of their kids. Cyclists ought to move in various ways."[13] That is, we apply standards to people all the time even when they are not able to do any better. Feldman refers to these standards as "role oughts."[14] Importantly, these prescriptions seem true and appropriate even when the agents referred to would never be able to fulfill those demands—even if they did the best they could. A teacher may be incompetent, a parent incompetent, and a cyclist uncoordinated but, in virtue of their respective roles, there are standards that apply to them. Feldman grants that we do not hold them to standards at superhuman levels and that our prescriptions are, in some way, limited by their capacities. However, these standards are not limited to the point of having voluntary control over one's behavior or outcome and may still outstrip the capacities of particular individuals.

Feldman argues that we should think of our epistemic oughts as role-oughts. These oughts apply simply because, as cognitive agents, we frequently take on the role of believers.[15] In fact, we involuntarily "take on" the role of believer but, according to Feldman, this does not undermine the appropriateness of the epistemic prescriptions. We are also involuntarily thrust into the role of eaters and breathers for which we can rightly say there are ways one ought to eat and breathe. For any activity that we take on, voluntarily or involuntarily, there is a standard of performance (consistent with the goal of the role) that applies.

Feldman purposefully sets his sights short. He admits that "some terms, especially those associated with *praise* and *blame,* are to be reserved for voluntary behavior."[16] His argument is limited to our use of requirements and permissions (that is, our use of the word 'ought'). He rightly points out that Alston's actual attack is against the notion of being "free from blame for believing"[17] but dismisses the idea because he does not think that "this more narrowly defined notion of deontologism is so natural or common."[18]

Whether Feldman has shown that we normally break the ought-implies-can principle, Feldman runs into trouble because role oughts only seem to apply in virtue of taking on a role. Kornblith and Chris-

man object on the grounds that doxastic oughts seem to be *categorical;* that is, they seem to be appropriate and applicable even when we are not actively taking on a role (or no matter which role we are taking on). This may be true for certain moral beliefs: It seems true that I ought to form a belief about whether what I am doing is wrong even if I am not engaging with whether or not it is. We can also appeal to negative prescriptions. It seems true that, no matter what I am doing, I should never form the belief that I am an alien, a robot, a princess, or so on. Kornblith and Chrisman also point out that it would be true from a role-ought perspective that, should a kleptomaniac involuntarily take up the role of thief, it would still be true that she should steal.[19] However, *categorically,* the opposite is true: No matter whether she adopts a certain role (or not), it seems true that she should not steal.[20]

Kornblith offers an alternative approach. He argues that Feldman is right that some oughts are derived from an evaluation of good performance (and that they can reject the ought-implies-can principle) but, instead of evaluation of roles and their goals, they derive from human ideals. He agrees with Feldman that there is a large middle ground in between prescriptions beyond *anyone's* capacities and those beyond merely a few individuals'. However, he claims, because ideals apply to individuals regardless of whatever role they are taking on (as a thief, slave, etc.), they maintain a normative force that 'role-oughts' lack, namely, categoricity. Similar to Feldman, Kornblith's treatment of epistemic oughts is orthogonal to the question of voluntary control. For instance, in developing his view, he considers some of our epistemic oughts deriving from our natural mental capacities. That is, he envisions certain 'oughts' about what kind of information we should believe (such as 'if one sees rain, one ought to believe it is raining') based on how effectively we can make valid inferences (e.g. the reliability of our visual-belief system). However, even if we grant that we should be sensitive to the reliability, or capacity, of our mental systems, it still doesn't have any bearing about whether exercising those mental capacities is considered voluntary. The debate about ought-implies-can, and arguing for oughts/ideals placed in some middle ground, does not yet engage with our original problem of epistemic responsibility or blame. This problem is more clearly seen in an ac-

count of epistemic deontologism provided by Matthew Chrisman.[21]

Kornblith and Feldman have developed accounts based on treating our epistemic ideals as applying to our *actions*. Chrisman argues that ideals, insofar as they apply to beliefs, operate a certain way based on the fact that beliefs are *states*. Chrisman argues that ideals, when they apply directly to actions, should be action guiding. This action guiding requirement, argues Chrisman, is especially prominent in our moral ideals: moral ideals that fail to lead to action or engage with how we behave seem to be bad ideals. Chrisman worries that because beliefs are involuntary, any epistemic ideals, the way Kornblith understands them, will fail to engage our behavior. Instead, he points to many kinds of ideals that do not apply to actions, but to states. For instance, our doctor may tell us that we ought to have a certain blood pressure or a mechanic might say that a bike ought to have a certain center of gravity. These oughts can be true, like the kinds of oughts proposed by Feldman and Kornblith, even when our blood pressure or the balance of a bike are not within our (or someone else's) voluntary control—that is, nobody can choose to do anything to change it.

Chrisman, in developing his account, follows a distinction used by Wilfred Sellars between "ought-to-do"s and "ought-to-be"s.[22] That is, oughts that govern *actions* are 'ought-to-do's, or **rules of action**, and oughts that govern *states* or, *ways of being* are 'ought-to-be's, or **rules of criticism.** Rules of action presuppose the 'ought-implies-can' principle, whereas rules of criticism do not. Rules of criticism do not have to directly engage with our behavior. Because of this, Chrisman believes, we can maintain the categoricity of epistemic evaluations desired by Kornblith, but avoid the worry of non-volunteerism, because they apply to states instead of actions. However, according to Chrisman and Sellars, rules of criticism do indirectly engage with our behavior; they imply rules of action.

Chrisman highlights a few ways that rules of criticism imply rules of action: conditionally, universally, and existentially. Take the following rule of criticism: X ought to be φ. According to the *conditional view*, the rule of action that is derived is:

> If someone is responsible for X's being φ, then that person ought to

> do what he/she can (ceteris paribus) to bring it about that X is φ.

According to the *universal view*:

> Everyone ought to do what he/she can (ceteris paribus) to bring it about that X is φ.

And, finally, according the *existential view*:

> Someone ought to do what he/she can (ceteris paribus) to bring it about that X is φ.

Ultimately, Chrisman does not endorse one of these methods over any other and leaves it an open question whether some or all of them are appropriate methods for deriving rules of action.[23]

The problem discussed above in Feldman and Kornblith (and now Chrisman) becomes prominent in Chrisman's analysis: Chrisman relies on the notion of responsibility in explicating this view. This is significant for a number of reasons. Firstly, it highlights the fact that the kinds of analyses being offered by Feldman, Kornblith, and Chrisman are inappropriate (or missing the goal of) developing a view of how we can, if we can, be responsible for our beliefs. Secondly, it highlights a need for an account of responsibility because, whether or not people can be responsible for their own beliefs or the beliefs of others will determine how effectively we can derive rules-of-action from rules of criticism. If Chrisman fails to be able to derive rules of actions from rules of criticism, then he loses the supposed advantage he had over Kornblith concerning an analysis of human ideals.

From our critique of volunteerism above, we have reason to be suspicious both of whether we are responsible for own beliefs (which is the original charge) and also of whether we can effectively modify other people's beliefs. The reasons we lack control over our own beliefs seem to apply to any scenario in which we were attempting to control or change someone else's beliefs. That is, while we can do our best to provide evidential reasons for a certain belief, we cannot control or effectively influence how they appraise or respond to that evidence

(just like we cannot control how we respond/appraise evidence we are evaluating). Given this, it seems unlikely that the *universal* or *existential* methods of deriving rules-of-action would adhere either.

All three of these theories fail to address Alston's primary challenge against epistemic responsibility and blame. Some, like Feldman's, purposefully sidestep it or, as we saw, dismiss it as being unnatural or uncommon. Kornblith and Chrisman do not engage with the issue or, as we saw with Chrisman's account, even attempt to build a theory presupposing some aspect of responsibility.

This failure mirrors an important question about moral oughts and responsibility. We might think that moral oughts are derived from human ideals and yet wonder if people are ever responsible for failing, or succeeding, their mandates. For example, before we reproach a teacher for failing to be an ideal teacher, we do not *merely* wonder if she had any sort of contractual obligation and whether or not she met the requirements. We inquire into how much effort she exerted, whether she knew she was doing a poor job, and whether she had other courses of actions available to her (like finding another job). Once we have some understanding, we begin to form a judgment about the character of the teacher. If we find out that she is a terrible teacher because she just doesn't care, we will form negative judgments about her above and beyond the initial judgment that she is a bad teacher.

Before attempting to develop an account of epistemic blame and responsibility, it will helpful to review Alston's challenge against it.

Responsibility without Effective Voluntary Control

One of the specific claims Alston was arguing against was the claim that we had control over our beliefs (and, therefore, epistemic responsibilities) because we had the ability to deliberate when forming a judgment. One of his opponents, Roderick Chisholm, aptly describes the intuition motivating this claim:

> When a man deliberates and comes finally to a conclusion, his decision is as much within his control as is any other deed we attribute to him. If his conclusion was unreasonable, a conclusion he should not

> have accepted, we may plead with him: "But you needn't have supposed that so-and-so was true. Why didn't you take account of these other facts?" We assume that his decision is one he could have avoided and that, had he only chosen to do so, he could have made a more reasonable inference. Or, if his conclusion is not the result of a deliberate inference, we may say, "But if you had only stopped to think", implying that, had he chosen, he could have stopped to think. We suppose, as we do whenever we apply our ethical or moral predicates, that there was something else the agent could have done instead.[24]

Alston's response to Chisholm is to show that, when we attempt to form specific proposals about how voluntary control figures into the practice of deliberation, they all turn out false. For instance, Chisholm could be suggesting that if we deliberate, we are able to gather new information, or uncover new evidential reasons, from which we may be able to *will* a judgment—essentially, exhibit immediate voluntary control over our beliefs. However, obtaining reasons, be they evidential or practical, does not allow us any control over which of those reasons will apply in the outcome of the belief. Our beliefs will still be outside our will, no matter how many new reasons we obtain to believe something. We can imagine that I may deliberate and call to mind many practical reasons why I should believe a certain proposition without gaining any ability to act on them.

Instead of relying on an implicit thesis of immediate voluntary control, Chisholm could be suggesting that such deliberation offers us "long range" control over our beliefs. This long range control be analogous to "indirect" voluntary control, discussed near the beginning of the essay. This is a more plausible, but ultimately misguided, thesis. This is because deliberation, as we've already discussed, is limited to the activity of mobilizing, or discovering, certain reasons (evidential or otherwise). While we can choose to engage a certain question, and engage in deliberation, all we are *merely* doing voluntarily is deciding to inquire into some proposition. Alston writes:

> Claims like [Chisholm's] . . . ignore the difference between doing A in order to bring about E, for some definite E, and doing A so that

> some effect within a certain range will ensue . . . At most it shows that I have long-range voluntary control over whether I take up *some* propositional attitude toward *some* proposition (Alston 2005, p. 71)

According to Alston, what we need in order to display effective voluntary control and, therefore, to be responsible for our beliefs/judgments is the ability to take up a *specific* attitude toward a *specific* proposition. Which, as discussed earlier in the essay, is something we cannot do.

Despite Alston's critique, the intuition that we somehow have responsibility over (some) of our beliefs because of some fact about deliberation (or our ability to deliberate) still seems right to me. This is for the specific reason that deliberation does not merely entail the ability to take up some attitude toward a proposition but, rather, the ability to take up some propositional attitude *of a certain quality* toward some proposition. It would be meaningless if the only quality that could be applied to our beliefs after deliberation was the mere quality of having been extensively deliberated about. However, deliberation can give our beliefs many kinds of qualities, such as being evidentially rich, one-sided, narrow-minded, and so on, depending on how we deliberate.

Our deliberation (or lack thereof) will (at least some of the time) play a determining or necessary role[25] in the eventual judgment we form. When we do have voluntary control over something that plays a determining role in a state then we can, sometimes, be held responsible (or blameworthy) for that state even when we do not have direct voluntary control over it. Alston introduces this option as an option for the epistemic deontologist:

> Consider the general point that we can be blamed for a state of affairs F, provided something we voluntarily did (didn't do) and should have not done (done) was a necessary condition (in the circumstances) of the realization of F. That is, F would not have obtained had we done (not done) something we should have done (not done). If my cholesterol buildup would have been prevented had I regulated my diet in the way I should have done, but didn't, I can be blamed for that buildup, whether or not I have direct effective voluntary control

> of my cholesterol level. (2005, p. 74)

That is, occasionally the only reason a belief we have formed is defective (that is, not true) is specifically due to our activity of deliberation. The belief was only defective because we did not deliberate and, had we deliberated, would not have been defective. Returning to the grocery store example from the beginning of the paper, I could have engaged with the question of whether buying non-organic food was appropriate and searched my memory for information or evidence. If I had done that I would have swiftly remembered the conversation I had had with my partner and formed the true belief that buying non-organic food is inappropriate. My failure to deliberate was the only thing preventing my belief having a negative epistemic quality, namely, the quality of being false. Just as I would be blameworthy for my bad cholesterol level if my diet was really the only thing preventing its eventual deterioration, I would be blameworthy for my false belief.

The number of things we can voluntarily do that influence beliefs extends beyond deliberation. We can search for evidence, talk to our friends, consult experts, and so on. Alston is aware that our beliefs are affected in these ways; however, he claims these activities as constituting intellectual, instead of epistemic, obligations. He claims that:

> S is intellectually to blame for believing that p *iff* if S had fulfilled all her intellectual obligations, then S's access to relevant considerations, or S's belief-forming habits or tendencies, would have changed in such a way that S would not have believed that p. (Alston 2005, p. 77)

One of my intellectual obligations might be to search my memory for possible evidence of a proposition's falsity. If doing that would have given my belief-forming mechanisms or habits access to information that would have resulted in a different belief, then I am blameworthy.

Although Alston grants that we have intellectual duties, he maintains that they are not properly considered epistemic. His claim that they are not epistemic is strange since, as he himself claims, our intellectual responsibilities are "rooted in an obligation to seek the true and avoid the false in belief" (Alston 2005, p. 76) which strikes me as

eminently epistemic in value. He is claiming that we have a responsibility to seek the truth and avoid falsity but it is not an epistemic responsibility. Given that epistemologists define the domain of epistemology as that governing truth, excluding our intellectual obligations from epistemic value seems inappropriate.

How, then, does Alston reconcile this oddity? Alston claims that we actually do not every have epistemic responsibilities because any apparent ones (such as the "intellectual" ones) do not actually yield epistemic value.

Alston and Epistemic Value

Like many epistemologists, Alston considers the domain of the epistemic, both in normative and descriptive theorizing, to pertain processes or goals related to 'truth' (as against being concerned, e.g., with 'the good' as morality is). He writes,

> [Epistemic Desiderata] are those features of beliefs or bodies thereof are valuable from the epistemic point of view, defined in terms of the aim at acquiring true rather than false beliefs about matters that are of interest or importance to us. (2005, p. 47)

BonJour provides a similar sentiment:

> If our standards of epistemic justification are appropriately chosen, bringing it about that our beliefs are epistemically justified will also tend to bring it about that they are true. If epistemic justification were not conducive to truth in this way… then epistemic justification would be irrelevant to our main cognitive goal and of dubious worth. (2005, p. 7-8)

The problem with the proposed model of intellectual obligations is that it does not sufficiently connect us (or, our beliefs/judgments) with the truth:

> It is prima facie conceivable that being formed in a way that does

> not depend on violations of intellectual obligations should be a way of rendering a belief probably true. Nevertheless . . . there are very many sorts of cases in which one does as much as could be reasonably expected of one in the way of voluntary acts leading up to a given belief without the belief's thereby acquiring any considerable likelihood of truth. (Alston 2005, p. 78)

The analogous case for a potential moral obligation is establishing whether fulfilling that obligation yields something good. His argument is as follows:

> (1) The fulfillment of our intellectual obligations be considered epistemically desirable iff it is the case that it entails a "considerable likelihood of truth."
> (2) The fulfillment of our intellectual obligations does not entail a "considerable likelihood of truth."
> (C) The fulfillment of our intellectual obligations is not epistemically desirable.

Before critiquing this argument, it is worthwhile to explore some other criteria that are epistemically desirable to understand why our intellectual obligations are restricted to truth-conduciveness. Besides truth-conduciveness, Alston acknowledges that something can also be epistemically desirable if it is "favorable to the discrimination and formation of true beliefs" (Alston, 2005, p. 43). Epistemically desirable features include:

> (A) Having some high-grade cognitive access to the evidence, and so on, for a belief (B).
> (B) Having higher-level knowledge, or a well-grounded belief, that B has a certain positive epistemic status, or
> (C) That one can carry out a successful defense of the probability of truth for B.

As opposed to *rendering* a belief likely to be true, these qualities already presuppose the truth of a belief. These constitute higher-level

knowledge that indirectly leads to more true beliefs. That is, if I know that certain kinds of beliefs, like visual beliefs, are generally truth-conducive (or that a belief is true because it was a visual belief) I will be in a position later to form true beliefs about what I see. However, fulfilling our intellectual obligations does not yield these kinds of higher-level epistemic resources.

The last group of Alston's epistemic desiderata includes features of *systems* of beliefs that we typically hold as goals of cognition. These include having beliefs that are coherent, provide reinforcing explanations, result in an agent acquiring some kind of *understanding*, and so on. Importantly, these desiderata are not valuable because they *produce* true beliefs. They are valuable only when they rely on true beliefs. Alston claims that "unless truth can be assumed, these features of belief systems would fail to exhibit the intrinsic cognitive desirability that would otherwise attach to them" (p. 46). Alston's argument for including them on the list of *epistemically* desirable qualities is their otherwise intrinsic cognitive value and their relation to a system of true-beliefs. Our intellectual obligations fail to qualify for this category, according to Alston, because even if we established some intrinsic value independent of truth, "that value is in no way dependent on being associated with a preponderance of true beliefs" (2005, p. 78).

Returning to Alston's primary argument, he admits that, *prima facie*, it seems like premise 2, the claim that fulfilling our intellectual obligations actually does yield a high-probability of true-beliefs is true (see the quote above). However, Alston points to scenarios in which it clear that not only have we fulfilled our intellectual obligations, but that belief still is not likely to be true. For instance, it is a common occurrence that people are too unintelligent to form the correct belief. No amount of deliberation will help a first-year student grasp the finer points of Descartes if they simply cannot understand it. In contrast, something like a "sufficiently reliable belief-forming process" is, by stipulation, truth-conducive so long as it is fulfilled. Alston also points to areas of our lives where it would be unreasonable to hold any such obligations even though our belief-forming habits are notoriously unreliable. For instance, he claims accepting information from authorities is unreliable and, at the same time, unreasonable to keep people

obligated to constantly check sources, follow-up on the information, and so on. Alston also points to frequent occurrences of when people are overcome by certain beliefs or certain ideas, like God. No deliberation (or other voluntary action) will make them attune to evidence or reasons that will yield a different (true) belief. I will return to this part of Alston's argument later. For now I will concern myself primarily with the notion of epistemic value.

The first problem with Alston's argument is that the requirement of "a considerable likelihood of truth" for epistemic value proves too much. There are many things I would take to qualify as an epistemic ideal (and on the basis of some conduciveness criterion) that would not qualify under Alston's thesis. Take, for instance, the ability to reliably form accurate analyses of the validity of syllogisms. Those who can reliably do it are more ideal epistemic agents than those who are unreliable. The problem is that such ability does not guarantee any sort of reliability in the belief about the conclusion of the syllogism. This is for the simple reason that, as valid or invalid as many syllogisms are, so many of them are simply not sound. The premises rest on false conclusions; therefore, the vast majority of syllogisms are false, and our supposed ideal is rendered epistemically impotent. We might extend this to ideals about Bayesian belief updating. That is, it seems plausible to me that having our beliefs undergo perfect Bayesian updating (that is, the total incorporation of prior, likelihoods, and so on into our posteriors) is epistemically ideal. However, so many of our beliefs are false in so many domains that, even with perfect belief updating in light of evidence, our beliefs are not guaranteed likely to be true.

These two examples highlight an important class of epistemic activities not captured by Alston's taxonomy. Many of the constraints of rationality (with Bayesian inference being a controversial epitome) are, at least with regard to intuition, epistemically valuable. This is not because they guarantee a certain likelihood of truth but because they are the most sure-fire way of getting at truth. As such, they are heavily implicated in the *aim* of truth even if they do not guarantee its success (or a certain threshold likeliness of success). Their value then, is not an objective threshold (like the truth-conduciveness quality Alston defends) but a conditional truth-conduciveness. This leads to a strange

picture. The fulfillment of these different rational or evidence-seeking obligations makes us ideal *conduits* for truth. Given the right environment (this includes aspects of our cognition outside our will) our *will* maximizes the likelihood of true belief. Alston is right to point out that many other parts of us fall prey to problems like sheer unintelligence or defunct belief-forming mechanisms. This leads the surprising disconnect between our otherwise great epistemic characters and our lousy beliefs. That is, we can do the very best we can to follow the aim of truth without thereby rendering it likely that our beliefs are true.

I will not argue for a conception of epistemology that is not tied to the acquisition of true beliefs.[26] Instead, I propose a more lenient criterion of when epistemic obligations can be induced. Instead of limiting epistemic desirability to only beliefs (or processes which guarantee a significant likelihood of truth), I propose that epistemic value is added to a process (or activity, function, etc.) whenever it makes the likelihood of truth in the resulting belief more likely. Now, instead of only having intellectual obligations which always guarantee a significant likelihood of belief (of which there are none), we have intellectual obligations whenever fulfilling them leaves our resulting belief more likely to be true than had we not.

This more lenient criterion has the benefit of applying to many of qualities of rationality that intuitively seemed to provide epistemic value. Additionally, many of the activities we envisioned as constituting our intellectual responsibilities qualify. Things like searching for more evidence, spending more time on an argument, and so on, make our resulting beliefs at least a little bit more reliable (even though there are times, as Alston has apparently pointed out, when they are not very effective). By accepting a more lenient version of the conduciveness criterion we are most likely limiting the role that epistemic blamelessness can play in other normative matters. BonJour and Alston's emphasis on a robust connection or guarantee of truth is a popular notion and has played an important role in developing accounts of justification and knowledge. Given this, the epistemic deontologist may have to accept that being epistemically responsible or blameless does not entail being epistemically justified or having knowledge. If I am right about the importance that epistemic responsibility and blamelessness play,

in our social lives at least, then this fact is not disconcerting.

The account we have developed here posits the following: We have the epistemic goal of acquiring a true belief. From this epistemic goal, we have epistemic obligations to do what we can to increase the likelihood of acquiring a true belief. These obligations supervene only on behaviors we have voluntary control over, severely limiting the extent to which we actually do have responsibility (or can be blamed) for a belief. However, when we fail to fulfill those obligations we are epistemically irresponsible (and blameworthy) for our belief being less likely to be true than it could be). When our irresponsible epistemic behavior is (at least significantly) implicated in the production of a bad belief (false, ill-grounded, etc.), we are blameworthy for that belief.

We can also understand how, when other kinds of obligations in other domains of value invoke the goal of truth they, in fact, are implicating epistemic value. When we have a moral duty to form true moral beliefs and we act epistemically irresponsible, we are rightly thought of as both *immoral* and a *bad epistemic agent*. We are then persons who are insufficient concerned with both our ethical duties and with truth. The failing of our intellectual responsibilities (which we now understand as epistemic responsibilities) can yield blame in more than one domain of value.

Whether or not our appeal to the more lenient value criterion is successful, it is important to engage with Alston's second premise. There is at least one reason we should hope it is not true: mainly, if it is, that then those kinds of intellectual obligations would rarely, if ever, operate. If, even after all of our hard work, we are still incredibly unreliable, then those responsibilities are not appropriate for the goal of achieving true or false beliefs. Recall that such responsibilities are frequently co-opted by moral and practical obligations on the basis of *acquiring true beliefs* being relevant to those domains. It seems as though one crucial reason we have a moral obligation to discover the correct moral principles (or form the correct moral appraisals of actions and so on) is so that they increase our ability and tendency to do good. If it turns out that actually, all of the different ways we *could* have (indirectly) influenced our beliefs never helped achieve that goal then the goal of those moral obligations is lost and we lose those de-

rived obligations (epistemic or not). The same problem applies in other domains of value.

For reasons I will get to in a moment, I do not find Alston's evidence for Premise 2 to be very compelling. However, work in recent evolutionary and social psychology has mounted a challenge to the idea that our deliberation and active reasoning play even a positive role in our goal for true beliefs. Psychologists, responding to years of research showing that certain 'failures' in reasoning are in common and robust in human beings, have proposed and defended a model that claims that our tendency for rational error is built in to the very function of explicit reasoning (see, for instance, Mercier & Sperber, 2011). That is, our explicit deliberation and reasoning function to reinforce previously held beliefs and reject evidence against them. Instead of abiding by principles of rationality that increase the likelihood of getting our beliefs right, deliberation and explicit reasoning actually make us more dogmatic and closed-minded.[27]

We can conceptualize the challenge as follows. If we leave our mental obligations too lenient, then they are easy to accomplish but they are susceptible to the biases we mentioned above. Alternatively, we can attempt to limit the effect of the biases by making our mental obligations specific (and balanced against these biases). For instance, we might say that we have an obligation not merely to seek out evidence but, rather, to specifically seek out evidence contrary to our beliefs. However, these are more difficult to fulfill, even when a person is trying to do so. If a person fails even they were attempting to fulfill their obligation they are not blameworthy. On the one hand, we have a people succeeding their obligations but those obligations themselves not increasing the likelihood of truth. On the other, our obligations are well-formed but people are blamelessly failing to accomplish them. Either way, these mental obligations are ill-suited for epistemic value.

I focus on deliberation and explicit reasoning for two reasons. First, many of the obligations we were supposing actually applied involved things like "searching our memory for evidence," "analyzing arguments," and so on. If it turns out that these are subject to the bias mentioned above, and that engaging in them does not actually help us get to the truth, then many of our supposed obligations disappear. The

significance of this point is related to the second reason. Recall from Chisholm's long quotation above that many of the intuitions about the role we play in our epistemic lives is connected to our mental behavior. One of the supposed advantages of establishing a theory of epistemic deontology was its ability to capture many of these kinds of intuitions. So, although we could probably find other kinds of epistemic obligations (that deal more with non-mental behavior), if we lose the obligations over our mental lives, the victory is bittersweet.

The solution to this problem is to realize that we have not done a good job specifying what our mental obligations are and in what kinds of contexts they apply. While it is true that we cannot have obligations such as "update your beliefs according to Bayesian ideals" because people are generally not cognitively capable of that, we can highlight, in certain contexts, what kinds of mental activities tend to be the most conducive. So, when forming beliefs about subjects we are not familiar with, our epistemic goal is to defer to authority. When Alston gives the example of the student who is unable to understand Descartes, we can simple alter the obligation. No longer does that student have an epistemic obligation to deliberate to personally mull over the philosophical arguments but, rather, to ask his professor or T.A. or so on. Even if he *still* could not form the right beliefs, at least approaching authority in this case has made a true belief more likely to result.

We can respond to the challenge in part by admitting that sometimes we should not rely on our mental prowess alone. However, there are still plenty of times when what our obligation consists in is remembering certain pieces of relevant information. Remembering specific facts is, in plenty of relevant scenarios, not limited by the argumentative biases noted above. Recall the aforementioned example about forming the wrong belief because of my failure to remember what my partner said to me mere minutes earlier. Such obligations as searching for memory can greatly increase the likelihood of forming a true belief in these kinds of scenarios. Finally, we might consider that some people have far more intellectual obligations than others. A brilliant scientist is going to be held far more responsible for their beliefs than someone vastly undereducated. The epistemic deontologist can happily accept this development.

Alston's challenge about the reliability of our voluntary behaviors, and the arguments related to that, are serious worries. I have attempted to show that perhaps there is hope for finding the right kinds of rules and obligations that will tip the scale of truth-conduciveness in its favor. However, the challenge is fundamentally an empirical one and is dependent on the deontologist actually creating these systems of responsibilities, which I do not have the space to do here.

Conclusion

As we saw, Feldman, Kornblith, and Chrisman were primarily interested in the legitimacy of claims such as "you ought not believe x." We came to favor an account based on epistemic ideals, with the various obligations falling in a large "middle ground" what people were capable of and what they were not. The middle ground is a balance between rejecting the "ought-implies-can" principle and the notion of engaging with, and guiding, our behavior. For instance, if a person is presented with a syllogism and they form the wrong belief about the validity of that syllogism, we might rightly say that they ought not have done that because it is within the human ideal to correctly interpret a syllogism, even if it is just beyond that particular individual's capabilities.

Where all of these theories fell short, however, was in providing an explanation about when people were responsible for those failures and when they were not. It seemed that, because we have no will to believe and, therefore, no voluntary control over our beliefs, that we are *never* responsible for our beliefs. However, we found that we can be blamed for our beliefs when we fail any number of epistemic responsibilities we hold over our voluntary actions *and* those are significantly responsible for a defect in our belief. One challenge we dealt with was that of understanding how fulfilling those obligations was properly understood as epistemic in value.

As long as those responsibilities are epistemic, we have the tools to connect the projects of Feldman, Chrisman, and Kornblith, with an account of epistemic responsibility. Consider another case of someone being presented with a syllogism. We maintain our epistemic "ought"

that they should believe that the syllogism is false. This person, however, forms the belief that it is in fact a valid syllogism but they do so because they didn't actually read the entire thing (merely the first premise, say). In this case, they violated our epistemic ought *and* they are responsible and blameworthy for it.

Although we have established that we have epistemic responsibilities, and that we can be blamed for our bad beliefs, we must still understand that any kind of responsibility and blame we have for our beliefs is derivative of our responsibility for the parts of our epistemic lives we actually have control over. However, because we do have control over certain epistemic parts of our lives, we can be better or worse epistemic agents, much in the way we can be moral or immoral. People's expectations of the kinds of beliefs they have are informed by the way we lead our epistemic lives, that is, the voluntary parts we have control over, and the epistemic situations we find ourselves in. ■

I am grateful for the helpful comments suggested by the editors of Arché. *I have especially benefited from many stimulating conversations with and comments from Trevor Brothers, Geoffrey Sayre-McCord, and Maegan Fairchild.* - CC

Notes

1 William Alston, *Beyond "Justification": Dimensions of Epistemic Evaluation* (Ithaca: Cornell University Press 2005), 58-73.

2 We often make a distinction between when a person *deserves* blame or is *responsible for* an action or state and when it will serve some instrumental purpose or function to punish them or change them. For instance, the first time a child draws on the dining room walls with crayons we may be upset and punish them. Most likely, we will do this without making any sort of evaluation about what kind of child they are. However, if, the very next day, they go back and draw on the wall again, then we engage in a different kind of judgment. Although they did something *bad* in both circumstances, only in the latter would we say they are culpable or blameworthy for that bad behavior. When we judge in terms of blame or desert, the focus of our disapprobation moves from external features (like an individual's situation) to internal features (like a person's character).

3 Alston, *Beyond "Justification": Dimensions of Epistemic Evaluation*, 73-75.

4 Notice that this syllogism does not mention at all *epistemic responsibility*. If we can show that we do have *some* kind of responsibility over our beliefs then we still need to show that it is properly considered *epistemic.* This task will be taken up later in the essay.

5 Of course, I can imagine a crazy world or an unlikely set of experiences (in this world) which, if they occurred, I may form a belief this belief.

6 Alston, William. "The Deontological Conception of Epistemic Justification." *Philosophical Perspectives* 2 (1989): 129.

7 As noted by Feldman (2004), there are very rare circumstances in which we can control what we believe in virtue of controlling our environment. For instance, if I want to believe that the light is on I can go over to the switch and turn it on. However, as noted by Feldman, these examples are to disparate among the vast array of our beliefs to constitute an argument for belief volunteerism.

8 Bennett, Jonathan. "Why Is Belief Involuntary?" *Analysis* 50.2 (1990): 87-107.

9 Feldman (2004) gives the example of controlling one's belief about whether the light is on by simply controlling whether the light is on. He dismisses the example on the grounds that such influence over our beliefs is too rare to be relevant for arguing that we are responsible for our beliefs.

10 Feldman, "The Ethics of Belief," 667-695.

11 Kornblith, Hilary. "Epistemic Obligation and the Possibility of Internalism." In *Virtue Epistemology: Essays on Epistemic Virtue and Responsibility*, by eds. A. Fairweather and L. Zagzebski, 231-248. Oxford University Press, 2001.

12 Chrisman, Matthew. "Ought to believe." *Journal of Philosophy* 105.7 (2008): 346-370.

13 Feldman, "The Ethics of Belief," 676.

14 Using "should"/"ought" locutions when a person could not have done differently (or better) is not restricted to the domain of role oughts.

15 We might also consider the relevant role to be "a person trying to discover the truth" "a person attempting to deduce a conclusion" and so on. I think that these are wrapped up in the concept of a believer but I do not think that, should they be different, it matters to the dialectic of this essay.

16 Feldman, "The Ethics of Belief," 676

17 Specifically, epistemic justification as freedom from blame from believing but my target is not whether blame is connected at all with justification, but rather if it is properly applied at all.

18 Feldman, "The Ethics of Belief," 677.

19 This example is taken directly from Chrisman (2008).

20 Feldman may have recourse against some of these objections. First, it is not clear that we are ever not in the role of an epistemic agent. One might claim that evaluating forming judgments about our environment (and a whole host of other topics) is an automatic function of our psychology. It cannot be the case that we are actually uninterested in the truth because being so is a necessary function of our psychology and therefore always subject to such a performance analysis. Because roles are the kinds of things that can be ever present, Feldman may also respond to the case of the unfortunate kleptomaniac arguing that she is still in the role of being a moral citizen (or something akin to that).

21 Chrisman, "Ought to Believe," 346-370.

22 Sellars, Wilfred. "Language as Thought and as Communication." *Philosophy and Phenomological Research* 29.4 (1969): 506-527.

23 See Chrisman (2008), p. 12-14 for a fuller description of deriving rules of action from rules of criticism.

24 Chisholm, Roderick. "Lewis' Ethics of Belief." In *The Philosophy of C.I. Lewis*, by ed. P. Schilpp. Salle, IL: Open Court, 1968, 244.

25 We might also weaken it to simply playing a "major" role. This way we can find someone partially responsible, partially blameworthy, and so on.

26 For such an attempt, see Foley, Richard. *The Theory of Epistemic Rationality.* Cambridge, MA: Harvard University Press, 1987.

27 The mere fact of the frequency of these failures of cognition is enough of a challenge. The added fact that this may stem from *proper* function of cognition does not, on its own, mean we should or should not engage in deliberation.

Reconciling *Eros* and Neuroscience: *Maintaining Meaningful Expressions of Romantic Love in a Material World*

by ANDREW J. PELLITIERI*
Boston University

Many people currently working in the sciences of the mind believe terms such as "love" will soon be rendered philosophically obsolete. This belief results from a common assumption that such terms are irreconcilable with the naturalistic worldview that most modern scientists might require. Some philosophers reject the meaning of the terms, claiming that as science progresses words like 'love' and 'happiness' will be replaced completely by language that is more descriptive of the material phenomena taking place. This paper attempts to defend these meaningful concepts in philosophy of mind without appealing to concepts a materialist could not accept.

Philosophy engages the meaning of the word "love" in a myriad of complex discourses ranging from ancient musings on happiness, to modern work in the philosophy of mind. The eliminative and reductive forms of materialism threaten to reduce the importance of our everyday language and devalue the meaning we attach to words like "love," in the name of scientific progress. Faced with this threat, some philosophers, such as Owen Flanagan, have attempted to defend meaningful words and concepts important to the contemporary philosopher, while simultaneously promoting widespread acceptance of materialism. While I believe that the available work is useful, I think more needs to be said about the functional role of words like "love" in the script of progressing neuroscience, and further the important implications this yields for our current mode of practical reasoning.

I will begin by addressing the controversial claim of eliminativ-

* kierke@bu.edu. Received 1/2011, revised December 2011.

ism in its most radical form. The claim is that much of the content of our everyday language will soon be eliminated and displaced by the emerging scientific understandings coming from our advancements in neuroscience. I believe I will be able to show that the proponents of the eliminativist perspective need to reconsider some of the underlying assumptions driving their arguments. I will then turn to reductionism, which I understand to be the view that our ordinary terms such as "love" and "consciousness" will be subject to *intertheoretic reduction,* "where a new and very powerful theory turns out to entail a set of propositions and principles that mirror perfectly (or almost perfectly) the propositions and principles of some older theory or conceptual framework."[1] The reductionist perspective can be considered as a revisionary response to some of the more popular criticisms made against eliminativism. Though reductionism is a more distinct position than eliminativism, I will still argue that there is not enough evidence given by current scientific studies to warrant the promises of a reductive movement, and the implications from conceptual history and available neuroscientific research are not as damning as the reductionist project. I will conclude by offering an alternative framework to use when approaching progress made by science, wherein use of our everyday language fits comfortably with our understanding of materialism, i.e. the idea that everything, including mental phenomena, is composed of material, and as a result of physical interactions.

Materialism carries a lot of weight among modern scientists, as there is fear that any alternative requires dualism. Assuming the arguments against dualism to be strong, it is useful to keep this theory in mind when trying to understand the perspective of the following reductionist proponents (though attention should be paid to the false dichotomy: either reductionism or dualism).

The account I am trying to provide is also a rejection of dualism. The problem is conceiving of love in a way that fits in with the materialist world view that contemporary philosophy seems to call for. In the arguments I make for the role of *eros*, I must also account for the conception of any idea I utilize, including that of consciousness, within a materialistic framework. In order to provide a satisfactory reconciliation of all of these ideas with the materialist restrictions, I must

thus reconsider the following questions: Does love require a dualistic worldview? Does consciousness exist in the material world? And, how do we empirically account for the existence of love?

Eliminativism

In the philosophy of mind eliminativism is the position that conscious phenomena such as beliefs, desires, sensations, perceptions, will be proven not to exist as neuroscientific knowledge advances. Folk psychology should and will accordingly be abandoned in favor of neuroscience. This claim obviously yields deeply troublesome implications for our ordinary conception of love and other emotions. If eliminativism is correct, we should expect love not to exist as well. In this chapter I will examine the arguments for eliminativism and argue that none of them succeeds.

Churchland's Arguments

Paul Churchland has authored some of the most prominent eliminativist arguments. He claims (1984) that the resistance to eliminative sentiments is merely an attachment to a flawed folk psychology. He defines eliminative materialism as "the thesis that our commonsense conception of psychological phenomena constitutes a radically false theory . . . so fundamentally defective that [it] will eventually be displaced, rather than smoothly reduced, by completed neuroscience."[2]

His arguments embody the following ideas: Mental states known as emotions fail to cleanly correspond with complex neurological brain states. In our attempts to communicate our mental states, we employ an inefficient and ultimately problematic language. Once we have a universal understanding of the physical events that take place in accordance with the mental phenomena, we will much more effectively communicate our mental states in a descriptive neuroscience-based language. It is also thought that this point in progress will alleviate some of the problems of "other minds" and some of the descriptive misunderstandings rooted in folk psychology, such as "how memory works, or how we manage to retrieve relevant bits of information in-

stantly from the awesome mass we have stored."[3] The immediate appeal of such ideas would make up for the failure of our everyday forms of explanation to adequately account for mental illness, differences in intelligence, sensorimotor coordination, how the retinal image allows for perception of all three dimensions, and certain aspects of memory.

The two foundational assumptions that emerge from the eliminative arguments can be summarized as such: first, the belief that the type of descriptive warrant attributed to *eros* is less valuable than the type attributed to scientific description; and accordingly, the subsequent belief that the former descriptive warrant justifies the elimination of our current expressions, in terms of "folk psychology," for a new set, one defined solely in scientific terms. Churchland puts forth three claims to this effect:

> (1) Our early folk theories have traditionally, throughout conceptual history, been displaced entirely by more sophisticated theories; thus folk psychology will inevitably follow in this tradition of displacement.
> (2) The widespread explanatory, predictive, and manipulative failures of folk psychology will fold from the explanatory poverty and failure.
> (3) The *a priori* probability of eliminative materialism is substantially higher than that of either identity theory or functionalism.

Charles Siewert (1998) adequately captures my initial reaction to these claims when he writes: "everyday mind talk and brain science are supposed to be incompatible; but to this we must ask: 'Why?'"[4] Churchland gives two reasons at the beginning of his work: First, it seems very unlikely that the arrival of an adequate materialist theory would be able to directly correlate and align with our concepts of folk psychology with our concepts of theoretical neuroscience. He goes on to argue further that "because our common-sense psychological framework is a false and radically misleading conception of the causes of human behavior . . . it is an outright *mis*interpretation of our internal states."[5] Both of these claims depend on the validity of the three

outlined arguments, to which I will now turn my attention.

Argument #1: Warrant from Conceptual History

Churchland (1988) begins his first argument for eliminativism by calling attention to the claim that, "as the identity theorist can point to historical cases of successful inter-theoretic reduction, so the eliminative materialist can point to historical cases of the outright elimination of the ontology of an older theory in favor of the ontology of a new and superior theory."[6] Consider his example of how "witches" were eliminated from our ontology when the theory of psychosis was introduced. Psychosis is a relatively common affliction among humans and "in earlier centuries its victims were standardly seen as cases of demonic possession, as instances of Satan's spirit itself, glaring malevolently out at us from behind the victims' eyes."[7] Some weaker examples he gives, like the belief in the "noble soul," are more easily eliminated because they are non-observable phenomena, but witches were at one point empirically accounted for, and yet the conclusion was made that "the concept of a witch is an element in a conceptual framework that misrepresents so badly the phenomena to which it was standardly applied that literal application of the notion should be permanently withdrawn."[8]

But how well do the implications of this example extend to the existence of love? Churchland grants that, since we are working within a materialistic framework, "only empirical research can tell us where on [the reductive spectrum] our own case will fall."[9] In the witch example, the empirical observation being described by the word "witch" was that of individuals "engaged in incoherent, paranoid, or even murderous behavior."[10] The word "witch" thus carried with it many meanings: an individual with the ability to do magic, an individual with power from the devil, an individual who exhibited incoherent, murderous, or paranoid behavior (hereafter referred to as IMP-behavior). We have not eliminated any of these meanings, as far as I can tell, from our use of the word "witch" when it is still used in reference to supernatural fantasy; instead we have eliminated the object to which those meanings refer. Thus, the elimination, it seems, took place in that our attributions of demonic possession and the quality of being-a-witch,

Figure 1. The left side of the diagram lists an empirical observation [EO] and the various understandings available [UN] *prior* to some scientific advancement [SA]. The right-hand side lists the understandings *after* said SA. The numbers rank possible importance of the understanding.

were eliminated from our empirical understanding of accounts of IMP-behavior; when Churchland wants to eliminate the terms in our mental language, he is denying that they refer to anything real. The elimination that took place in this example can thus be illustrated as in Figure 1 (which will help when considering the future possibility of extending the example to mental terms).

The "witch understanding"—that IMP-behavior was a characteristic of a witch—was easily displaced by the scientific understanding because it didn't seem to have any empirically recognizable meaning or understanding that functionally correlated with the event being described. Upon first consideration, it seems unlikely that this will extend to something as complex as the observations made about love. To examine whether or not the model given by Figure 1 will extend to the contents of folk psychology I will proceed to identify the empirical observation that folk psychology is attempting to describe, tease out the available understandings, and analyze the eliminative or reductive warrant of current scientific advancements. In order to adequately accomplish these three tasks, I must first critically address Churchland's second argument that claims that folk psychology puts forth understandings that suffer from just such explanatory poverty as that of the witch understanding. I will thus continue by showing that he mischaracterizes our ability to currently account for conscious mental events.

Argument #2: The Explanatory Poverty of Folk Psychology

The first major problem with the account Churchland (1981) gives

of folk psychology is his unfair treatment of its empirical content; he frames the argument against reduction alongside an argument for dualism, characterizing the non-eliminativist as akin to the *dualist* who "expects that [folk psychology] will prove *ir*reducible to completed neuroscience, by dint of being a nonredundant description of an autonomous, nonphysical domain of natural phenomena."[11] Because we are trying to give empirical accounts of our mental states, we do not want to require a dualistic image of the mind where there is anything immaterial being described (and this image is not very popular among empirical scientists). The false dichotomy given between reductionism and dualism is exposed by many, including Owen Flanagan. It is thus useful to consider his method for evaluating meaningful concepts in the material world when arguing against what is widely accepted by most philosophers within the philosophy of mind as a mischaracterization of folk psychology.

Flanagan's Method

In his controversial attempt to reserve the ideal of 'meaningful happiness,' Owen Flanagan (2007) established a methodical construct for evaluating any meaningful term within the confines of materialism. Flanagan begins with the following question: "Is there anything substantive that can be said about how best to find meaning and to live purposefully, to achieve *eudaimonia*, given we are fully natural beings?"[12] Flanagan is specifically concerned with the question of ultimate happiness and fulfillment, employing the term *eudaimonia*. The philosophical psychology of eudaimonics is the all-encompassing attempt at unifying the other sciences to explore fulfillment and flourishing. The idea is that having a well-developed and progressive eudaimonic practice will lead to knowledge of how to guide practical reason, and expose how we ought to live. Wilfred Sellars claims "the aim of philosophy . . . is to understand how things in the broadest possible sense of the term hang together . . . to know one's way around . . . it is therefore 'the eye on the whole.'"[13] Eudaimonics would reach this goal for philosophical enterprises as it tries to reconcile all types of science with the naturalist disposition. It follows from the existence

of the social sciences that everything may be understood scientifically; Flanagan offers a proof of this possibility:

> (1) Humans are natural creatures who live in a natural world.
> (2) According to neo-Darwinian consensus, humans are animals.
> (3) Human practices are natural phenomena.
> (4) Art, science, ethics, religion, and politics are human practices.
> (5) The natural sciences and the human sciences can, in principle, describe and explain human nature and human practices.
> (6) Therefore, sciences can explain, in principle, the nature and the function of art, science, ethics, religion, and politics.[14]

There is room here to recognize some confusing assumptions that might be made. First, by understanding the human practices one might need to trace their causal antecedents and consequences. As the human sciences are included in "the sciences" it should be recognized that there are other evaluative measures. It does not require that the intentional objects, those things under evaluation (art, ethics, etc.) be reduced to "mere things." This assumption would turn the world into a mere collection of scientific objects, a reductionist movement. The mistake lies in the lack of a distinction between two claims: "the claim that science can explain everything we think, say, and do—that it can provide a causal account of being; [and] the claim that everything can be explained scientifically."[15] The latter claim provides an account reduced to a collection of scientific idioms; it is here that the philosopher's vocation requires us to hold onto meaning and not fall victim to this reductionism, if possible, in order to live meaningfully.

Flanagan examines the scientific image in terms of common psychology, including the proper and improper conclusions that follow from this image; there are three important confusions to note surrounding the scientific image: 1) the view of scientism; 2) the view of the individual and its natural characteristics; and, 3) the view of consciousness and its normative features. The first and third confusions are the most important confusions for our examination, so I will skip the second as it is more useful in Flanagan's conception of freedom and causality, and not as useful as the view of consciousness. Carefully

considered, "the scientific image . . . need not be reductive, eliminativist, or disenchanting."[16]

Scientism

Patricia Churchland (2011) evades the overreaching claims of scientism—the claim that science can and will explain everything, and do everything—by qualifying the belief that the scientific approach to understanding morality does not threaten the arts or humanities; yet, she still claims that, "it is true that philosophical claims about the nature of things, such as moral intuition, are vulnerable. Here, philosophy and science are working the same ground and evidence should trump armchair reflection."[17] Though the second part of her claim alludes to the retention of some reductive sentiments, the first part makes the concession of the existence of spaces of meaning beyond scientific expression. In other words, there are meaningful expressions that exist beyond understanding in scientific idiom. Considering accounts of art, music, and other imaginative enterprises helps illuminate why scientism fails. The *nature and function* of music, for example, can be causally accounted for, its physical manifestations can be subject to physics and mathematic examination, its lyrics seen in terms of cause traced by the artists functionality; the explanation of these elements do not identify "the production itself as something that can be demonstrated scientifically."[18] The meaning of the love song is understood and it does use words, but the idiom is not a scientific one. The song "Brown Eyed Girl" may cause you to understand the love felt for the girl, but it does not give causal support for the feeling. Causal statements about the physiology of perception of the song abound, but these statements cannot approach the meaning of the song—that is what requires creativity over rationality, linguistic tricks over propositions, and transcendent understanding over scientific explication. The finding is that not everything worth expressing is scientifically expressible, *not* that some expressions are beyond scientific understanding. The materialist must allow other meaningful expressions outside of science, especially those that identify their meaning with their physical correlate. I will now examine conscious experiences and

attribution of mental states in this way, now free of the assumption of dualistic requirements.

Consciousness

Charles Siewert (1998) has done influential work providing for the existence of consciousness in the face of the eliminativist proposition. Siewert argues that for folk psychology, in its attempts to explain human behavior, to be proven radically false on empirical grounds, the eliminativists suggest the idea that "the explanations we offer in terms of attitudes and experience are to be seen not only as constituting a theory, but as constituting a theory that is *inferior* to that which neuroscience does or will provide" (50). Further, because the inferiority has to be so strong as to warrant the rejection of our everyday idiom, the eliminativist must describe the behavior that is to be explained without using any of the terms to be eliminated. The requirements of the eliminativist perspective when considered in this way seem very difficult to meet, thus Siewert goes on to ask, "isn't it clear that our warrant for attributions of attitudes and experiences depends on how well these explain things? And isn't it clear that what these explain, if anything, is our observable behavior?"[19]

Even Churchland (1981) concedes that "folk psychology does enjoy a substantial amount of explanatory and predictive success, and what better grounds than this for confidence in the integrity of its categories?"[20] When we attribute typical attitudes like embarrassment or surprise to our experience, the descriptive force of the attributions seems quite powerful. If we attribute an attitude of confusion to action such as Sandra raising her hand we might be able to understand the motivation of asking a question. If, however, upon calling on her it turned out that her arm was having an involuntary spasm, it is not likely that such false attribution would warrant my rejection of all of my attributions of confusion to all of the cases where the act of raising one's hand successfully communicated confusion and the desire to ask a question. Also, through this example, we can notice that there is often more than one correct explanation of some episode of behavior. Sandra raising her hand could be described: scientifically using language

describing the set of muscles employed; folk psychologically using my attribution of confusion to her action; sociologically in terms of the cultural expression that the act signifies. The significant conclusion that Siewert (1998) would draw out of this example is the realization that, "'mere movements' can also be explained in the idiom of attitude and experience . . . our warrant for attributions of attitude and experience to others is not to be assessed entirely on the basis that it best explains what is conceived of in a thoroughly dementalized fashion."[21]

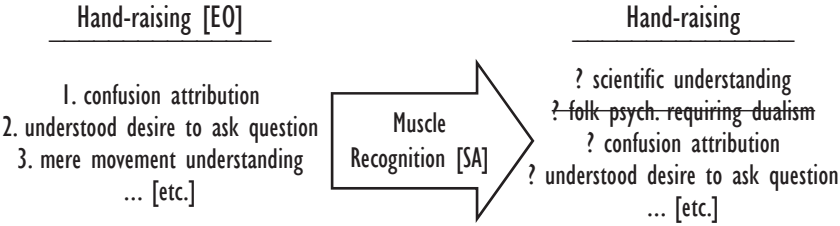
Churchland (1988) responds to the argument that eliminative materialism is false "because one's introspection reveals directly the existence of pains, beliefs, desires, fears, and so forth," by insisting that in both the case of the witch and that of familiar mental states, "precisely what is challenged is the integrity of the background conceptual frameworks in which the observation judgments are expressed."[22] Given the considerations made about the act of raising one's hand, we can illustrate, in Figure 2 (next page), the possibility of eliminating our understanding of our attribution of confusion, using the same model as Figure 1.

It is quite clear that, in many ways, the witch example does not extend to our everyday attribution of attitudes. The [SA] of muscle recognition might not warrant as much worry as the [SA] surrounding the much more confused case of *eros*. Thus, I will now turn to my understanding of these arguments, specifically when considered in terms of romantic love, in order to: 1) identify the empirical observation that *eros* is attempting to describe; 2) tease out the available understandings; and 3) analyze the eliminative or reductive warrant of current scientific advancements.

Romantic Love and the Witch Model

The empirical observation (EO) of *eros* takes many forms. In order to work *eros* into the model I have been using for the evaluation of the prospect of elimination, I will thus explore a wide range of possible EO, including the observations of: (1) John kissing Jane; (2) my first-person feeling of love toward individual S (attribution of attitude "love" to S); and (3) the businessman-who-loves-his-wife and is not

Figure 2. Empirical observation of Sandra raising her hand. The use of '?' instead of numbering displays the objective lack of importance of one understanding relative to another. The scientific advancement carries with it the presumption of materialism. However, this presumption does not eliminate any understanding previous to the SA; it merely anticipates that any understanding "requiring dualism" be eliminated. Note also that folk psychology requiring dualism is not included in the first column, as our actual folk psychology is not that confused.



cheating. In these explorations, I will tease out the various ways to describe and understand each observation, including the available neuroscientific advancements.

EO (1): "John Kisses Jane"

Though there are currently very few studies examining neurological activity during sex and other acts associated with romantic love, there is a lot of interesting, and relevant, information available about kissing. Generally, the findings of these studies are discussed in terms of evolutionary function, behavioral influence, or gender relations. One article that looks at these implications with specific regard to romantic love is "Romantic Love: An fMRI Study of a Neural Mechanism for Mate Choice" by Helen Fisher, Arthur Aron, and Lucy Brown (2005). The findings cited in this article can be used to frame a myriad of philosophical discussions concerning *eros* specifically:

> The range and variation of motivations and emotions associated with human romantic love are undoubtedly produced by many neural systems, acting in parallel and dynamic combinations. Nevertheless, several results support our hypotheses . . . that romantic love is

> primarily a neural system associated with motivation to acquire a reward, rather than a specific emotion; that this brain system is derived from mammalian precursors; and that it evolved as a mechanism to enable individuals to respond to sexually selected courtship traits and motivate individuals to make mate choice.[23]

With such advanced research methods, it will be easy to see how advances in our scientific understanding are still not threatening to our everyday language. If the eliminativist worry is valid, then why are we still using the word love within the very scientific studies that are supposed to provide its replacement?

The evolutionary understanding of kissing makes for an excellent point of departure. The biological function of mate selection required a system by which the organism could evaluate potential mates. The lips, as depicted on the sensory homunculus, are one of the areas of the body most densely with sensory receptors. The theory stands that kissing "conveys subconscious information about the genetic compatibility of a prospective mate [which is] consistent with the idea that kissing evolved as a courtship strategy."[24] Whatever the reason, the correlation between kissing and romantic love is taken to be evident, and the scientific understanding of how "romantic love" is evolutionarily functional is observed accordingly. In later sections, I will provide a more thorough account of the development of the surrounding evolutionary theory; but for now, the given summary provides enough information to consider the foundations for the belief that our scientific understanding of the events and actions correlated with *erotic* acts is continuously progressing.

It then seems intuitive that continuous progress yields inevitable understanding of all scientific descriptive correlates. These scientific descriptions are considered more accurate/valuable than those we currently use in our everyday talk of "love," they will progress past our folk conceptual understandings, requiring us to displace our use of "love," with the descriptive content defined in the findings of our new studies, i.e. the measured chemical reactions in the subcortical reward regions. But the meaning here does not seem to be reduced at all; instead, we just see an expansion of the scientific understanding. The possibility

of elimination seems even more distant as Chip Walter (2008) claims that even the current findings suggest that "the kiss continues to resist complete scientific discussion [as] close scrutiny of couples has illuminated new complexities woven throughout this simplest and most natural of acts."[25] Perhaps our current language is the best possible way to currently consider the studies since it is able to capture complexities otherwise inaccessible to a purely scientific description.

In her most recent work, *Braintrust,* Patricia Churchland (2011) also discusses the importance of love in mammalian evolution. She attempts to provide an account of morality that originates in the biology of the brain. The ability to tether ideas about "our nature" to something concrete in the world is realized through neuroscience; Churchland claims that, "Aristotle, Hume, and Darwin were right: we are social by nature. But what does that actually mean in terms of our *brains* and our *genes?* To make progress beyond the broad hunches about our nature, we need something solid to attach the claim to."[26] What neuroscience currently tells us about *eros* can be considered in terms of what neuroscience tells us about the evolution of value, morality, and social behavior. The previously cited article on kissing by Chip Walter (2008) revealed that "kissing unleashes a cocktail of chemicals that govern human stress, motivation, social bonding, and sexual stimulation . . . [and] one hormone, oxytocin, is directly involved in social bonding."[27] Oxytocin (OXT) is the most notable chemical associated with acts of *erotic* nature and attribution; OXT has the most notable appearance in social evolutions in the mammalian brain; OXT is also found in all vertebrates, but Churchland (2011) notes that "the evolution of the mammalian brain adapted oxytocin to new jobs in caring for offspring and eventually for wider forms of sociability."[28] The examination of *eros* is thus interwoven into the expression of social values. The prominent standing hypothesis posits the neurochemistry of attachment and bonding in mammals as the most important explanatory element. But Churchland goes on to ask the important question: how can neurons value something? Value statements seem to engage in one of the many spaces of meaning beyond simple scientific expression.

The complex interplay of OXT in attachment is seen to extend beyond the tight circle of offspring when it is put in terms of mate

attachment. In the 1970s, Sue Carter, a noted neuroendocrinologist, conducted some useful studies with prairie voles and montane voles. These studies revealed that the prairie voles have much greater OXT receptor densities than the montane voles in the *ventral pallidum* and the *nucleus accumbens* regions; both of these subcortical regions are identified with the reward-and punishment system. The prairie voles also demonstrated much stronger social behavior. They even displayed clear mate attachment and bonding after the first mating. Further, when experimenters blocked the OXT receptors, the treated voles did not bond after the first mating.[29]

These findings all indicate optimistic implications for the future strength of scientific description; however, the important question is whether these findings extend to *human* mate attachment, which appears to be more complex. Churchland cites anthropologists George Murdock and Suzanne Wilson, claiming that, despite apparent commonality of strong mating attachments, the flexibility of human mating arrangements can be seen in that "83% of societies allow polygynous patterns of marriage."[30] The recognition of patterns in the relation between the emergence of monogamy and non-biological events like the spread of agriculture in Eurasia, leads me to question the warrant of our current science, relative to the warrant of cultural memes or more general society-dependent moral practices.

This outline of the neurobiological explanation for the nature of our moral intuitions provides the means to understand the vast importance of expressions like the kiss, which evolved in such complex ways. Yet through the use of lay words like "trust" and "love" we can understand and discuss even the most complex evolutionary significance. The emphasis of the role for evolutionary understanding, along with the advancements made in our understanding of the material physiological response that takes place during a kiss, yields the illustration in Figure 3 of the possible reduction of the observation of John kissing Jane.

EO (2): My First-person "Feeling of Love" Toward Individual S

I will now consider the possibility of the importance of maintain-

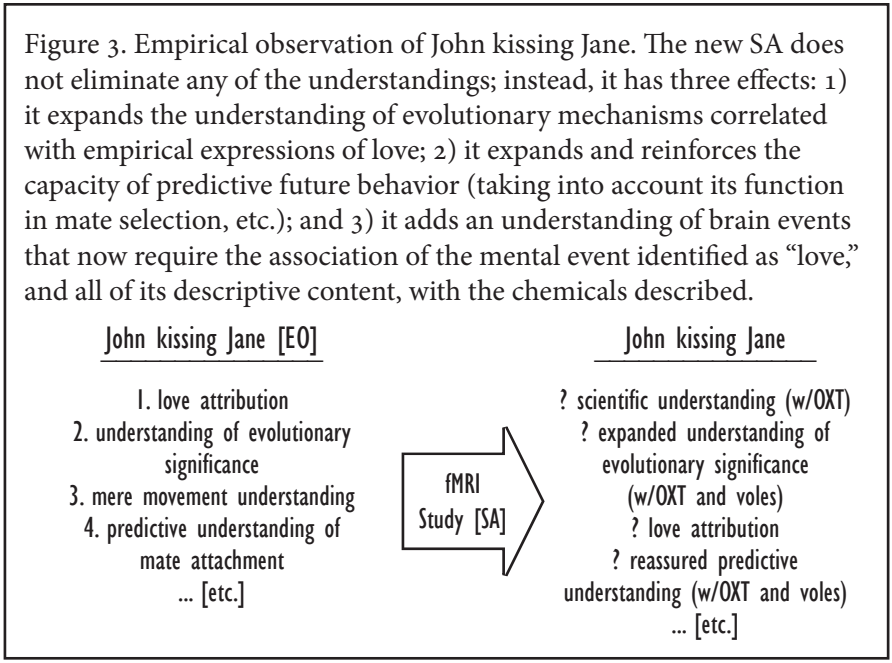
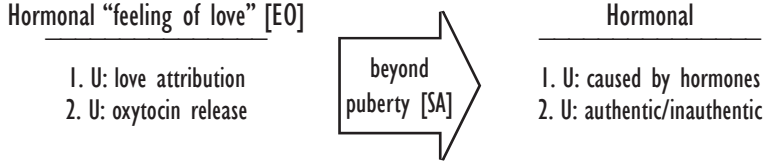


Figure 3. Empirical observation of John kissing Jane. The new SA does not eliminate any of the understandings; instead, it has three effects: 1) it expands the understanding of evolutionary mechanisms correlated with empirical expressions of love; 2) it expands and reinforces the capacity of predictive future behavior (taking into account its function in mate selection, etc.); and 3) it adds an understanding of brain events that now require the association of the mental event identified as "love," and all of its descriptive content, with the chemicals described.

ing all of my many understandings of my first-person attribution of the "feeling of love," by exposing the possibility for failure in multiple spaces of meaning. If during the first hormonal stages of puberty I believe I feel romantic love toward a female classmate, who does not warrant such strong attribution, I might come to realize upon later reflection that this was a false attribution. It would seem a scientific understanding, more adequately describing the causes and chemicals that created such a sensation, would have been more valuable, as it was definitively more accurate (this example captures the same explanatory failure as that concerned with understanding mental illness).

I am, however, later able to differentiate between my "artificial" hormonal experience (artificial in that it was caused by hormone imbalance, rather than the perceived beloved S), and a more "authentic" experience of the feeling of love (authentic in that it directly resulted as a response to the perceived beloved S). If we assume that the physiological response elicited in both the authentic and the inauthentic experiences would have been the same chemically, our scientific understanding might not be able to differentiate the significance; whereas,

Figure 4. Empirical observation of first-person "feeling of love."
"U" = "understanding of."



the understanding of the importance of the causality, will be meaningful for the understanding in the emotional space of meaning.

This example demonstrates the expanding importance of the meaning that we attribute to various mechanisms involving OXT, as the purely scientific understanding shows the possibility of confusing sexual love, with familial love, with platonic love, and so forth with most social bonding experiences. The illustration of this example in Figure 4 marks a shift in the authentic/inauthentic understanding, alongside the [SA].

EO (3): The Businessman Who Loves His Wife and Is Not Cheating

Consider the final example of a business man who is in Tokyo for a business trip for a month. Over the course of that month all of the physical neural events he experiences could be mapped out. On the last day if you asked him if he loved his wife back home it seems obvious that just because the records indicate that he was not experiencing any of the neurological correlates associated with "love" at any point during the business trip it does not mean that at any point he did not, in fact, still love his wife. This example of the failure of the possibility of mapping propositional attitudes leads up to the clear distinction that needs to be made between mental states and brain states, as mental states have a specific space-of-meaning that is important for attributing attitudes that wouldn't show up in chemical history.

Continuing this example, imagine the businessman encounters a woman who provokes the physiological responses that indicate sexual attraction and suggest mating behavior to follow. The physical under-

standing of the brain events does not adequately account for the man's love for his wife, which would, assuming he has some level of self-control, better predict his refusal to cheat with the woman.

Argument #3: Possible Trouble with Reduction

Now that I have given many of the arguments against eliminativism the final argument Churchland (1988) makes is easily dealt with. It is also helpful to discuss this argument immediately before I turn my focus to the prospect of reductionism, as the final claim concerning the *a priori* warrant of the eliminative perspective is set up against the warrant of functionalism and identity theory, which comes from reductionism. Churchland, in this vein, claims that "there are vastly many more ways of being an explanatorily successful neuroscience while *not* mirroring the structure of folk psychology."[31] This claim might be a reflection of the criticism of eliminativism, that he acknowledges the existence of, which claims that the major premise of eliminativism is only meaningful if it is the expression of a belief with an intention to communicate understanding, thus rendering its own claim incoherent if it succeeds. Churchland does not find the argument very productive.

A better criticism that Churchland also acknowledges exists within his work claims that eliminative materialism "is making mountains out of molehills. It exaggerates the defects of folk psychology and underplays its real successes . . . but the large-scale elimination forecast by the eliminative materialist is just an alarmist worry or a romantic enthusiasm."[32] Though it might not yet be apparent why we should not remain open to the possibility of reductionism, it has been my aim in this section to make it at least intelligible that our collective conceptual destiny is nowhere near the revolutionary end of the spectrum.

Reductionism

Paul Churchland (1988) claims that the central idea of reductive materialism is "simplicity itself: Mental states *are* physical states of the brain. That is, each type of mental state or process is *numerically*

identical with (is one and the very same thing as) some type of physical state or process within the brain or central nervous system."[33] The reductive promise of neuroscience is thought by some to be more warranted than the harsh eliminative promise. Paul and Patricia Churchland have even claimed that any characterization of their beliefs as 'harsh eliminativism' is misleading:

> [We] have no ideological stake in the revision being massive or minor, though our expectations lean toward the former . . . What we do believe is that our current framework is not sacred, that it is neither manifestly nor divinely given, and that 'obviousness' is a familiarity phenomenon rather than a measure of metaphysical truth.[34]

Regardless of whether we call their ideas "harsh eliminativism" or "revisionary materialism," as they would prefer, the unexamined assumptions at play remain the same. Though I would be remiss if I did not stop to make this clarification, I need to consider the radical cases of eliminativism and reductionism as some, like Paul Churchland (1984), might think they free themselves from criticism through qualifying their theory of reduction as "a possibility," that merely considers the intelligibility of the idea "that our collective conceptual destiny lies substantially toward the revolutionary end of the spectrum."[35]

Churchland points to four arguments for identity theory. I will argue that there is something wrong with the characterization of the introspection and its propositional content that is offered in support of these arguments. This is not meant to reject materialism, as Churchland posits as his worry,[36] but merely to expose that mental states are meaningful in such a way that makes it "ridiculous to expect a reduction from the behavioral level [and I will argue the conscious level] directly to the neuronal level."[37]

Four Arguments for Reductionism

The first argument reasons that the *purely physical origins* of each individual human (referring to the genetically programmed monocellular organization of molecules from which each person develops)

develop within a *purely physical system,* whose behavior arises from its internal operations and its interactions with the *physical world,* and thus "those behavior-controlling internal operations, are precisely what the neurosciences are about."[38] While neuroscience might offer useful information about the internal observations it has been shown in our previous discussion of scientism that materialism does not seem to warrant elimination alone.

The second argument given is similar to our earlier discussion of conceptual history, claiming that scientific explanation has been shown to be superior to other spaces of meaning previously throughout conceptual history. Again, this argument alone does not seem to warrant reductionism as our folk psychology "does enjoy a substantial amount of explanatory and predictive success. What better grounds than this for confidence in the integrity of its categories?"[39]

The third argument draws justification for reduction from the neural dependence of all known mental phenomena. We certainly could not have any mental states without any brain. The last argument alludes to the "growing success of the neurosciences in unraveling the nervous systems of many creatures and in explaining their behavioral capacities and deficits in terms of structures discovered."[40] But this does not warrant either what I have called scientism or what many philosophers call "category errors." I will turn to my argument of the meaningful property given to mental states that is not possessed by brain states. This property violates Leibniz' Law which states "that two items are numerically identical just in case any property had by either one of them is also had by the other."[41]

Argument Against Reductionism

Introspection is one of the ways to think about mental states and properties that appear starkly different from any neurophysiological states or properties. The reductionist posits that the four arguments above warrant the claim that "in discriminating red from blue, sweet from sour . . . our external senses are actually discriminating between subtle differences in intricate [neurophysiological] properties."[42] But through the example of the businessman-who-loves-his-wife, I have

shown that introspection might reveal and explain meaningful behavior in some instances where neuronal understanding fails. Thus, in some cases of predicting behavior, or understanding meaning of some behavior introspection seems to be a more useful tool then pure scientific understanding. Further, given Leibniz's Law of numerical identity, the following proof can be made:

(1) My mental states are knowable by introspection.
(2) My brain states are *not* knowable by introspection.

Therefore, by Leibniz' Law,

(3) My Mental states are not identical with my brain states.

One other property difference can be seen in that an empirical observation of whether something is sweet or sour is not the same as an empirical observation of whether a mental state was authentic romantic love or inauthentic romantic love (following the authentic/inauthentic distinction made in the example of attributing "feeling of love" to individual S).

Love and Progress: Living in Our Material World

The fear that results from the implications of eliminativists and the reductionists is that meaningful language, like the word "love," is hindering the progress of our current neuroscience. Churchland (1988) cites that more than adequate time has been allowed for the correction of the failures of folk psychology, and yet it "has enjoyed no significant changes or advances in well over 2,000 years despite its manifest failures."[43] I have yet to see an alternative theory step in, and until I do it does not seem productive to abandon or worry about the possibility of one coming along; it is only important that we understand that neuroscience might have a significant role in helping overcome some of the difficulties that folk psychology currently faces. The most productive way to move forward would be to recognize that both of these modes of study have a meaningful place and neither encroaches upon

the other. Folk psychology is not parasitic to neuroscience, but merely symbiotic to our progressive scientific system.

In our future explorations into the brain-basis for our values, our mental states, our emotions, et cetera, we should only take caution as to not forget the ways in which science informs other forms of understanding and vice versa. ■

Notes

1 Paul M. Churchland, *Matter and Consciousness*, (Cambridge: MIT Press, 1988), 26.
2 Ibid. 67.
3 Ibid. 46.
4 Charles Siewert, *The Significance of Consciousness*, (Princeton, NJ: Princeton University Press, 1998), 50.
5 Ibid. 43.
6 See Paul M. Churchland, *Matter and Consciousness*, 43.
7 Ibid.
8 Ibid. 44.
9 Ibid. 49.
10 Ibid. 44.
11 Paul M. Churchland, "Eliminative Materialism and the Propositional Attitudes," *Journal of Philosophy* Vol. 78: 1981), 72.
12 Owen Flanagan, *The Really Hard Problem*, (Cambridge: MIT Press, 2007), 61.
13 See Wilfred Sellars, op. cit., 5.
14 See Flanagan, op. cit., 21.
15 Ibid. 22.
16 Ibid. 36.
17 Patricia S. Churchland, *Braintrust*, (Princeton, NJ: Princeton University Press, 2011), 26.
18 Ibid. 23.
19 Siewert, op. cit., 54.
20 See Churchland, *Eliminative Materialism and the Propositional Attitudes,* 73.
21 See Siewert, op. cit., 56.
22 See Churchland, *Matter and Consciousness*, 47-8.
23 See Fisher op. cit., 60.
24 Chip Walter, "Affairs of the Lips," *Scientific American Mind*, Vol. 19, 2008, 26.

25 Ibid. 29.
26 See Churchland, P. S., *Braintrust*, 2.
27 See Walter, op. cit., 27.
28 See Churchland, P. S., *Braintrust*, 14.
29 "Curious about the effects of extra OXT, neuroscientists were surprised to observe that administering additional OXT in an otherwise normal female prairie vole results in a weakening of attachment to her mate" Ibid. 78.
30 Ibid. 57.
31 See Churchland, P. M., *Matter and Consciousness*, 47.
32 Ibid. 48.
33 Ibid. 26.
34 See Siewert, op. cit., 289.
35 See Churchland, P. M., *Matter and Consciousness*, 49.
36 Ibid. 33-34.
37 Patricia S. Churchland, *Brain-Wise: Studies in Neurophilosophy*, (Cambridge: MIT Press, 2002), 193.
38 See Churchland, P. M., *Matter and Consciousness*, 28.
39 See Churchland, P. M., "Eliminative Materialism and the Propositional Attitudes," 72.
40 Ibid. 29.
41 Ibid.
42 See Churchland, P. M., *Matter and Consciousness*, 30.
43 Ibid. 46.

About the Contributors

Alex Henderson is a senior at Georgetown University majoring in Philosophy and Government with a minor in Classics. He is particularly interested in Hellenistic philosophy and is in the process of writing a thesis on the influence of Epicureanism on the philosophies of Adam Smith and Karl Marx.

Michelle Mary Dyke is a senior at Yale University majoring in Philosophy. She is interested especially in early modern philosophy and epistemology, and plans to begin work on a Ph.D. in philosophy in the fall of 2012.

Corey Cusimano studied philosophy, psychology, and cognitive science at UNC-Chapel Hill. He is currently working at a language-processing lab at Brown University. He has philosophical interests in ethics, meta-ethics, philosophy of mind, epistemology, and experimental philosophy. In cognitive science, he is most interested in language processing, moral psychology, and decision theory.

Andrew Pellitieri graduated magna cum laude with a philosophy degree and with distinction from Boston University when he was nineteen years old. He is currently taking time off to write before he applies to doctoral programs in philosophy of mind. In his spare time, he makes money as a strategy coach in the *Starcraft* e-sports community giving lessons and participating in amateur professional tournaments.

Department of Philosophy

Faculty 2012-2013

Full-Time Faculty
Hugh Baxter
Alisa Bokulich
Peter Bokulich
David Bronstein
Tian Yu Cao
Daniel Dahlstrom
Juliet Floyd
Aaron Garrett
Charles Griswold
Jaakko Hintikka
Walter Hopp
Paul Katsafanas
Victor Kestenbaum
Manfred Kuehn
David Liebesman
David Lyons
Krzysztof Michalski
Robert Neville
Russell Powell
David Roochnik
James Schmidt
John Silber
C. Allen Speight
Susanne Sreedhar
Daniel Star
Judson Webb
Elie Wiesel

Adjunct Faculty
Alredo Ferrarin
Michael Grodin
Knud Haakonssen
Ghita Holstrom-Hintikka
Philip Ivanhoe
Simon Keller
Jürgen Renn
Steven Tigner

BOSTON
UNIVERSITY