The Idea

The Undergraduate Journal of Philosophy at Texas A&M University

Volume I Fall 2011

- Epistemic Normativity and the Egocentric Predicament pg. 1
  by Kody Dollins (Texas A&M University)
- Socrates: A Lyre in Tune pg. 15
  by Holland M. Kaplan (Texas A&M University)
- The Resolution of Incompatibilities in Kant's Compatibilism pg. 27
  by Jared Clifton (Texas A&M University)
- Suicide as a Response to the Realization of the Absurd pg. 38
  by Sarah Baver (St. Lawrence University)
- Burge vs. the Lawyers pg. 52
  by Anonymous

About the journal

The Idea is but one of other attempts at producing and sustaining an undergraduate journal of philosophy at Texas A&M University. Hopefully it can be noted from these numerous tries that maybe there is something valuable in producing a journal: to give students something not only to work collectively on, but a congealed form of their labor so that they and future Aggies can look upon it as a window to their past.

Dedication

Kant regarded the Idea, "as a goal... which itself always remains a kind of beyond," but Hegel took a step forward seeing, "that everything actual is only in so far as it possesses the Idea and expresses it." And just as Aletheia was but an, "opening of presence, [and] not yet truth," so are we here to take a step forward towards The Idea as, "the true as such."

It is only proper in this way to see that it will be up to future Aggie philosophers to determine whether or not The Idea will be able to, "return upon itself," anew each year. It is to them then, that we dedicate this journal.

Special thanks to

The Texas A&M Philosophy Club whose labor and money this journal could not have been without.

Editor-in-Chief
Paul C. Adamski

Assistant Editors
Jared Clifton
Jennings Kennady
Colin Noe
Reed Williams

Faculty Advisor
Dr. Scott Austin

For online archives, submission guidelines, or any questions, comments, or suggestions, please visit the journal at its website: <tamuphilosophy.wordpress.com>

EPISTEMIC NORMATIVITY AND THE EGOCENTRIC PREDICAMENT

Kody Dollins

Texas A&M University, 2013

Externalist accounts of justification claim an advantage over their internalist counterparts in at least one intuitively crucial respect: they are capable of addressing the question of whether or not a belief is ultimately and objectively true. However, the consequences of a prior commitment to the T-condition[1] of knowledge can manifest in short-sighted views on the J-condition. In response to this discrepancy, I will argue that placing external conditions on justification is inconsistent with common understandings of what it means to be justified in coming to a belief, to the extent that any theory failing to account for the agent's perspective threatens to render agency irrelevant and justification superfluous. In allowing for justifiers that are inaccessible, externalism undermines the normative component of justification intrinsic to our traditional understanding of epistemology. In place of external criterion on belief-justification, I offer an evidentialist account of justification structured upon norms of egocentric rationality.

Is it the believer or the belief that is justified? Perhaps one might say it is the belief, yet a belief apart from a believer is not possible, so that a "belief" (justified or unjustified) is inextricably linked to the believer whose cognitive action the belief is indebted for its very existence. Even if a belief is true, the believer may be unjustified in believing it if her reasons for doing so are based on inadequate grounds. So justification concerns more than the sheer truth of the matter, but involves something going on within the agent, some internal epistemic activity that is a necessary prerequisite for her to be considered justified. A person may hold a true

1. Truth-condition, a component of the traditional view that equates "knowing" with having a justified true belief:

A subject S knows a proposition p if

1) p is True;
2) S Believes that p is true;
3) S is Justified in believing that p is true.

I'll refer to these as the T-, B-, and J-conditions, respectively – and the "JTB" notion of knowledge, collectively.

belief (say, that a home in the neighborhood will burn down today), but for inadequate reasons (perhaps because every day this person believes a home will burn down in the neighborhood). The reasoning process behind that belief is what determines whether it is justified, regardless if today a home finally does go up in flames. We do not "justify" *beliefs* in evaluating whether they are true, because "justified" is a term ascribed to people and requires their having done well in rationally arriving at the beliefs they hold. When we judge a belief to be justified, we are rendering a verdict on the epistemic status of a believer.

"Egocentric predicament" is the term given to our confinement within our own minds and perceptions; it conjures images of that lonely Cartesian self, cut off from the external world in his own solipsistic universe. This subjective ingredient of the human condition invites questions of a distinctly normative sort, such as "What am I to believe?" "How do I do well in forming my beliefs?" These questions are as much at the heart of epistemology as epistemology is at the heart of the egocentric predicament, and at root of such self-reflection is the pursuit of truth, or a desire to *know*. However, as epistemologists, we must also not forget who it is that is seeking knowledge. Our aspiring know*er*, perhaps trapped in the travails of uncertainty, has no assurance that the result of his pursuit will be anything other than out and out skepticism. Our concept of epistemic justification should be helpful for him in his task; otherwise why should we suppose it is helpful for us? So to safeguard against justification becoming arbitrary or trivial to the one being judged, our theory must be mindful of the existential perspective of whom we are judging. After all it seems only natural that deliberations on *S*'s justificatory status require us to consider the subjective particularities of *S*'s cognitive state, experience, environmental influences, and so forth. Such considerations bear vital relevance to any adjudicatory talk of "how *S* should have done" epistemically. This is not to subjectivize truth,

or justification, as though either were defined merely by our experience of them (one can still be unjustified on an internalist schema). Rather, this is what *objectivizes* epistemic justification, by allowing for a sense of normativity capable of actually making sense to each and every one of us. Given the diversity of human experience and personality, it ought not to be assumed that the justification of humans will be any less dependent upon background information and context than is customarily required before you can say that you truly "know" someone on a personal level. Justification that has only truth as the impetus behind its normative force cannot address me on such a personal level. Let objective truth be the ideal goal – this need not divorce us from a commitment to the value of truth as being inseparable from its relation to (and meaningfulness for) us.

Alston has addressed the level-confusions that can arise in epistemology, which provides a distinction relevant to our discussion of epistemic justification. The two levels between which these confusions tend to occur are that of first-order *states* of being justified (irrespective of whether you are aware of or can articulate this justification) and higher-order *beliefs* about your own justification. To be justified only requires a certain sort of adequately-held belief about something or other; for instance, a young child's justified belief that he is in pain, despite being unable to describe the nature of the pain – only that it hurts. So to hold a belief that puts you in this initial "onto-epistemic" category of justification, you do not need to be able to state your justification, or even believe you are justified, let alone be justified in believing you are justified. This sort of self-contemplation and reflection on one's beliefs is more than is necessary to be justified in simply holding a belief. Still, it may be the case that onto-epistemic justification and awareness (conscious or subconscious) of *one's* own justification are frequently connected concepts in some way. Perceptual beliefs might be voluntary or involuntary, conscious or subconscious –

just not *un*conscious.

Externalism, as it relates to justification, can be broadly characterized as the view that what makes a belief "justified" is not *necessarily* something to which the believer always has access. It is perhaps best understood as a denial of its opposite, internalism, a position which insists on some kind of person-oriented understanding of justification. Where internalism sees it as essential that justification be applicable to the egocentric question of what we are to believe, externalism sees the descriptive truth of which beliefs have good-making features, and permits things to operate as good-making features (essentially, justifiers) of a belief that are not evaluable from the perspective of the believer – not that all justifiers must be external to the agent, but that they are not entirely internal either.

This view of justification places a heavy emphasis on the "truth" condition in the traditional analysis of knowledge as being "belief" which is both "justified" and "true." Whether or not a belief is actually true usually cannot be decided strictly from the believer's point of view, but is subject to external facts and conditions, facts which remain true or false independent of one's internal state and regardless of whether one has access to any evidence for or against these facts. Such a preeminent concern with the truth condition often leads to *reliabilism*, or the externalist's bottom-line: justification is fundamentally connected with truth, in such a way that unless our belief-forming processes are apt to lead us to beliefs that are *actually* true most of the time, we cannot be justified in assuming the reliability of our belief-forming processes, which by extension renders us unjustified in believing all of our beliefs.

The problem with the reliabilist view of justification is that it is still too focused on the "truth" condition, and allows this concern to bleed over into its conception of the "justified" condition. Perhaps looking ahead to the issue of knowledge, reliabilism makes

truth the focal point of epistemic justification, though because this is an unrealistic goal, is forced to settle for truth-conducivity as the deciding (if not only) criterion for justification. Since this is a concern with whether the belief-forming *process* is a reliable indicator of the truth of the belief, reliabilism is not altogether uninterested in the goings-on within the agent, but simply attaches a requirement to that internal epistemic process: that the justificatory status of its product (belief) be determined according to a standard that (unlike belief) does not find its origin in *S*. This requirement leaves the reliabilist thesis open to the charge of being unnecessarily reductionist, *vis-à-vis* a reduction of the justificatory efficacy of agency to inaccessible (and thus agent-neutralizing) external factors. Suppose reliabilism were guilty of its own sort of "level-confusion," by taking truth to be so decisively basic to justification, that justification becomes subordinate to truth. Perhaps we are justified on one level according to how well we have honestly reasoned *given what we had to go on*, and on another level, to be "justified" simply depends on how well our beliefs match up with the facts. The latter sense would justify something entirely different from the former, where one deals strictly with beliefs and the other with believers. For another illustration of this difference, suppose that *S* is a clairvoyant who has never heard of clairvoyance[2], and has no reasons for believing that the random true propositions that pop into her head from time to time are actually true. She simply believes them, for no evidential reason. Yet *S*'s clairvoyant episodes yield beliefs that are true every time. These beliefs are formed via a reliable cognitive process, which satisfies the reliabilist criterion for justification by producing beliefs which are actually true, yet it is clear that something is still

2. Laurence BonJour offers this thought experiment, along with various nuances, in "Externalist Theories of Empirical Knowledge," *Midwest Studies in Philosophy*, 5 (1980): 53-73.

amiss, if we are to call *S* justified. We are missing something having to do with *S*'s part, some work or act to be accomplished on the agent's end, so that it is clear that things like palpable cognitive negligence, irresponsibility, or epistemic luck can throw a wrench in justification. This speaks to the person-oriented nature of justification.

In review: who (as opposed to "what") is justified? It is the person or subject, *S*, engaging in the belief or acceptance of some proposition, *p*. What justifies *S*? On the internalist picture, it is her reasoning, in carefully coming to the formation of any firmly-held belief only after reflecting on what positive reasons are available for accepting this belief, what evidence may exist against this belief, other counter-beliefs which may be more rational to hold, and so forth, and still arriving at belief *p* despite having undergone this sort of conscious and honest reflection on the matter. While it may not guarantee the truth of *p*, it does preserve the rationality of *S*. Guarantees amount to certainty, and that is a Cartesian goal, which – at least for most items of knowledge – has long since been abandoned by epistemologists.

Now suppose we allow for the possibility that there are things external to *S* which function as justifiers of *S*'s belief that *p*. These external justifiers are beyond the scope of what is accessible for *S*, such that it would be impossible for *S* to cite one as justification for her belief, and still maintain her rationality. She has no access to them. Externalism argues that they can still factor in to whether or not S is justified because external factors ultimately determine whether or not many of our beliefs are actually true. The objective truth or falsity of beliefs about the external world is not decided by factors internal to the agent.

Still, if what justifies you can be something of which you have no comprehension, nor any possible access to whatsoever, then it seems tempting to want to do away with the "justified" condition in the appraisal of knowledge as "justified true belief," and sim-

ply define knowledge as "true belief." The justification condition seems intuitively motivated by the need to ensure that *S* has some proper grounds or basis for the belief; that there is some epistemic middle ground between sensory-inputs and belief-outputs that is crucial to justification and in need of being accounted for. Here – in line with the post-Gettier trend – I propose a thought experiment: suppose there were a machine, modeled after and capable of human-like thought processes, which was programmed with a unique memory of historical facts, and also programmed to have perceptive faculties and sensors that experience external stimuli in a very person-like way. Suppose further that it was exposed to such stimuli as the external world typically presents to a common human being, and on the basis of this stimuli and internal "reflection" (data processing, according to algorithms that appropriately mimic human cognitive tendencies), was capable of deductive and inductive inference, and generated various ordinary beliefs, along with some typical amount of unordinary beliefs. There is no particular superhuman ability in this machine; it was not designed to benefit humanity or help us solve any problems. The sole goal of its design was to make standard, generic beliefs about ordinary, arbitrary things in a way as typical of humans as possible. Would you care – apart from its general significance to advances in artificial intelligence – whether such beliefs were true or false? Would you consider this machine worthy of epistemic consideration, such that we might inquire as to its justificatory status? Surely there are involuntary mechanisms involved in our own processes of belief-formation, from our sensory equipment and perhaps even extending to less obvious cognitive faculties, but is it relevant to question whether a purely mechanistic process can be justified?

We would not call a robot or a computer program that, through human design, had been programmed to make probabilistic inferences about the future based on the past, "justified" in the

statements it asserted. One day soon robots may be capable of such artificial intelligence as to make human-like inductive inferences on the basis of autonomic experiential data.[3] The fascinating thing about the prospects of such technological advancement is that no matter how advanced, rational, or human-like the technology gets, necessarily it is still always numbers being manipulated by humans that enables the program. Even programs that seem self-aware in the sense of making real-time decisions based on what data is being recognized are reliant upon some formula programmed in by humans. Imagine if DARPA unveiled technology capable of predicting in advance when and where a crime has a strong likelihood of occurring, allowing law enforcement to be waiting. No matter how useful such a machine became, it would not qualify as a candidate for justification in the person-oriented sense. It would only qualify for justification on a true-false schema of reliability, and to any extent that it was "justified" in the reasoning it employed in arriving at beliefs that were true, this justificatory status properly belongs to those who coded the math and logic behind the program's "reasoning."

Searle famously argued that syntax is insufficient for semantics.[4] Computers may implement programs by processing strings of formal symbols and interpreting them according to their syntactical form, but we shouldn't say the computer "understands" a programming language in the way your mind is understanding the meaning of this sentence. Searle formulates this concern in his Chinese Room thought experiment, where he imagines himself playing the role of a CPU in a computer. Locked in a room, he's

3. Hans Moravec, a leading scholar on artificial intelligence at the Robotics Institute of Carnegie Mellon University, argues in a 2008 issue of *Scientific American* that by the year 2050, robot "brains," powered by computers capable of executing 100 trillion instructions per second, will begin to rival human intelligence.

4. John Searle, "Minds, Brains, and Programs," *Behavioral and Brain Sciences* 3 (1980): 417-24.

given batches of Chinese symbols that are semantically meaningless to him, along with a set of rules (written in English) for correlating certain sets of Chinese symbols with other sets of Chinese symbols. When he's given new scripts of Chinese text through a slot in the door, he's instructed to give back the corresponding symbols according to his list of rules. He can distinguish them only by their shape and sequence, but with practice he becomes efficient at associating certain shapes with their affiliated shapes, and when prompted he quickly outputs strings of Chinese text as though he were a fluent speaker. To Chinese speakers outside the room, his answers are perfectly intelligible. Perhaps the one who gave him the English instructions is a programmer giving a demo of her latest project to a Chinese audience, and from their perspective, Searle's been delivering insightful answers to a host of complex questions they've been asking of him. Yet he's understood none of this exchange, and knows no more Chinese than when he started. He's merely been performing computational operations on formal symbols, and still doesn't know the meaning of those symbols any more than a computer "knows" it's manipulating 0's and 1's.

The situation for actual machines is even less like the Chinese Room, because as a volitional agent Searle might've refused to play along, and instead begun pounding on the door demanding his freedom. This is a key difference between people and machines: whereas Searle could've chosen to tear the English instruction manual to shreds, in a computer that list of rules would've been the actual coding of the program, and binding on the computer's behavior in terms of how it's going to try to respond. Systems may crash because of conflicts between programs, but this is because the computer's dutifully trying to carry out two or more conflicting sets of instructions, rather than protesting its plight or having anything like a human emotion or mental state. Inorganic machines might appear to simulate human mental pro-

cesses, from an observer's perspective, but their operations are determined by their code, and they have no volition (from which it follows that they can't even desire to have volition). This simulation would not amount to a duplication of human mental processes. Even if we develop algorithms that enable machines to behave exactly like they're experiencing human emotion, it will be the code that determines this behavior, and renders their intelligence artificial. Perhaps the day will come when we can create organic machines that can make their own decisions based on their own emotions, ambitions, joys, fears; that get nervous around a pretty girl and embarrassed when they stick their foot in their mouth, and get why jokes are funny, and know what it is to find something "rewarding" or "fulfilling." You might argue that such a machine would be able to have a justificatory status – and you'd be right, since such organic machines are basically what we are. Consequently, perhaps such machines will never come, since it seems subject to all the ethical concerns associated with human cloning. This brings more clarity to our central point, for what bioethicist has misgivings about replicating a computer chip? For now and the foreseeable future, machines are not "alive" or "cognitive" in the same sense as human beings. There is an intrinsic value to being human that is necessary to even be eligible to have an epistemic status – or at least to *have* one in a deeper sense than merely being right or wrong.

The opposing sides of this debate reflect two very different stances on what should be the chief concern of our theory of justification. What is taken to be the goal of justification has a profound influence on the foundational beliefs one forms early on, beliefs that will likely have a strong influence on what further beliefs are cultivated upon further study. This suggests two structurally different concepts of justification, one being built around epistemic responsibility and egocentric rationality, and the other prioritizing truth and the reliability of our means for attaining

it. Reliabilism may be more amiable to our prospects for gaining knowledge, yet this comes with the burden of being a truth-oriented theory. It may modify things and say that the reliability must work through the agent's perspective, in that for justification to follow I must not possess subjective reasons for doubting the reliability of the process leading to the belief, but this still leaves me wondering which of my processes *are* reliable? Which ones *do* lead to truth? If I cannot know these things, why is reliabilism preferable to rationality? Especially if I *can* know whether my beliefs are rational, according to egocentric norms of such taken from experience and reflection. The burden is on me for being genuine and honest in such reflection, but at least this leaves me some measure of control over my justificatory status. This seems preferable to approaching epistemic justification from a "God's-eye" perspective that fundamentally gauges one's epistemic standing against an impersonal and often inaccessible standard, even if certain concessions are made.

While reliabilism represents a truth-oriented perspective with the guiding intuition being that truth is fundamental to justification, evidentialism addresses the question of justification to an agent-oriented understanding of rationality. If knowledge is attainable, and cannot be attained through epistemic luck or irresponsibility, then the vindication of *S*'s egocentric rationality is required before *S* can be declared justified with respect to knowledge anyway. Furthermore, even granting that truth and justification are both necessary in order for knowledge to occur, it does not thereby follow that true beliefs and (rationally) justified beliefs are of the same value for us.

Justification deals principally with people and is fundamentally based on their individual perspective. Again the contrary view that justification is belief-oriented and deals primarily with whether beliefs correspond to the objective truth of the matter, faces the issue of understanding how these beliefs can be justified

independently of the believer in whom they originate. Agents are justified according to how they should have done given their situation as a whole, in all its contextual complexity. In a sense the justification condition exists *because* of the reality of different perspectives, environments, patterns of reasoning, and value judgments. We determine how justified someone is in a belief by meeting them where they are at, and assessing their belief through their perspective, with the evidence available to them. We hold them accountable for ensuring that there is a proper relation between their beliefs and their evidence. This evidential relation grounds their belief in some adequate or appropriate basis for thinking it to be true.

To abandon an understanding of justification that is based on this evidential relation is to turn justification into a descriptive enterprise. It can still be evaluative, but this evaluation lacks the prescriptive force needed to help us with the egocentric predicament. In not meeting us where we are at, it can hardly be said to be "normative" for us in any relevant sense. Normativity requires applicability; that is the old maxim, "ought implies can." Since justification requires applicability to be relevant to our lives, and normativity is what makes justification applicable, *epistemic normativity* requires applicability to our lives. 'Justification' is the only distinctly *epistemic* (or person-evaluating, where "belief" would be person-*describing*) component of the JTB notion of knowledge. Kim gives a nice summation of this point:

> Neither belief nor truth is a specifically epistemic notion: belief is a psychological concept and truth a semantical-metaphysical one. These concepts may have an implicit epistemological dimension, but if they do, it is likely to be through their involvement with essentially normative epistemic notions like justification, evidence, and rationality. Moreover, justification is what makes knowledge itself a normative concept. On the surface at least, neither truth nor belief

is normative or evaluative. . . (383)

The evidential relation is a non-negotiable aspect of not only justification, but the whole enterprise of traditional epistemology. Kim is addressing issues with considering Quine's "naturalized epistemology" a form of epistemology at all, yet to the extent that externalist criterion of justification are allowed to factor into or influence our understanding of agent justification, externalism can be seen as fair game for similar criticisms. So far as normativity is at risk, justification (and with it, epistemology) faces equal danger.

References

[1] Alston, William P. 1985. "Concepts of Epistemic Justification" *The Monist*, 68: 57-89.

[2] Alston, William P. 1980. "Level-Confusions in Epistemology" *Midwest Studies in Philosophy*, 5: 135-150.

[3] Bonjour, Laurence. 1980. "Externalist Theories of Empirical Knowledge" *Midwest Studies in Philosophy* 5: 53-73.

[4] Foley, Richard. 1990. "Skepticism and Rationality" *Doubting: Contemporary Perspectives on Skepticism* (Dordrecht: Kluwer Academic Publishers): 69-81.

[5] Kim, Jaegwon. 1988. "What Is "Naturalized Epistemology?" *Philosophical Perspectives, V.2: Epistemology*: 381-405.

[6] Searle, John. 1980. "Minds, Brains, and Programs" *Behavioral and Brain Sciences* 3: 417-24.

SOCRATES: A LYRE IN TUNE - AN EXPLORATION OF THE PERCEIVED INCONSISTENCY BETWEEN PLATO'S CRITO AND APOLOGY

Holland M. Kaplan

Texas A&M University, 2013

Socrates, a prominent Athenian philosopher, was sentenced to execution by hemlock in 399 B.C. for the charges of impiety and corrupting the youth. Socrates' defense against these allegations is outlined in Plato's dialogue, the *Apology*. In Plato's *Crito*, we witness Socrates' friend, Crito, present justifications for why Socrates should escape from prison. Several apparent contradictions arise between these two works. In the *Apology*, Socrates appears to perceive the law as something that need not always be followed, while in the *Crito* he seems to regard the law as an unwavering bulwark against which we cannot act. It seems that Socrates thinks that he can disobey the law given the circumstances in the *Apology*, but still abide by the sentence agreed upon by the jury. He appears to endorse civil disobedience in the *Apology*, while sanctioning universal obedience of the law in the *Crito*. Delving into the literature provides many controversial interpretations of this ostensible inconsistency. While George Grote, a British historian from the mid-eighteenth century, argues vehemently that the *Crito* and the *Apology* are diametrically opposed, we can discern consistency by working through other authors' analyses of Socrates' conceptions of his duty towards justice, the law, and God, his perceived stubbornness to not stop philosophizing, the personified speech of the Laws, and finally the controversial persuade-or-obey doctrine.

Between the *Crito* and the *Apology* Socrates claims independently that his loyalties belong to justice, the law, and God. Socrates' various anecdotes during his trial generate contradictions. In the *Apology*, Socrates claims he will follow the laws of the state to preserve his honor, but asserts that his first duty is to God:

> I should indeed have wrought a fearful thing. . . if when the commanders you chose stationed me at Potidaea and Amphipolis and Delium, I remained where I was stationed, as others did, and ran the risk of death; but when it was God who stationed me, as I thought and believed, obliging me

> to live in the pursuit of wisdom, examining myself and others – if then, at that point, through fear of death or any other thing, I left my post. That would have been dreadful indeed. . . (28d-29a)

In the *Crito*, Socrates declares that by living in the city, he has entered into an implicit agreement to exhibit allegiance to the state. Socrates says in the *Crito's* personified speech of the Laws, "It is not holy to use force against a mother or father; and it is far more unholy to use force against your country." (51b-c) However, in the story of Socrates and the Thirty Tyrants, Socrates refuses a direct order to retrieve Leon from Salamis for execution, thus blatantly refusing to obey what he perceives as an unjust and unholy command (*Apology* 32c-d), forcing us to determine whether we have come upon an inconsistency in Socrates' perceptions of justice. Though Socrates complies with his commander while at war (and directly through his commander, the state) and he abides by God's order to philosophize, in the tale of the Thirty Tyrants Socrates unabashedly disobeys the orders given to him. This act of disobedience can be viewed as Socrates placing his allegiance with justice, since the actions of the Thirty Tyrants were technically legal, but still unjust.

Reginald Allen, in his book *Socrates and Legal Obligation*, addresses Socrates' scattered objects of loyalty by claiming that in the *Crito* and the *Apology* Socrates is consistent in his view that obedience to all law is not required. Allen identifies that demands requiring unjust deeds, such as the orders of the Thirty Tyrants, act outside the law and, therefore, do not mandate obedience. In such a situation, to obey the law would be akin to committing an act of impiety against God. Conversely, while at war, Socrates receives orders given by the legitimate leadership of his commander and thus has to obey them. We can deduce from Allen's assertions that he views Socrates as being loyal to justice in all cases. As Roslyn Weiss points out in her book, *Socrates Dissatisfied*, Socrates

equates justice with God. Therefore the view of both Allen and Weiss is that Socrates is, above all else, loyal to justice, and that through justice he is loyal to God. In his book *Socrates: Philosophy in Plato's Early Dialogues*, Gerasimos Santas also advocates consistency between the two texts, but for a different reason than Allen and Weiss. Santas centers his argument on the idea that the two dialogues address different types of disobedience. Socrates' disobedience in the *Apology* regarding the Thirty Tyrants is public and can be classified as civil disobedience. In the *Apology*, Socrates is attempting to obey God. Conversely, in the *Crito*, escape from prison by Crito's request would have been done in a sly manner. By escaping from prison Socrates would be surreptitiously disobeying both God and the will of the laws (Santas 50). Conversely, A.D. Woozley argues in *Law and Obedience: The Arguments of Plato's Crito* that the views on Socrates' loyalty in the two texts cannot be reconciled. He says that the *Crito* requires citizens to obey all laws, even if the laws are unjust. Woozley distinguishes between three of Socrates' claims: (1) A man must obey the law unless it is unjust, (2) a man must obey the law unless he believes it to be unjust, and (3) a man must obey the law unless he persuades "them" it is unjust. Woozley believes that these claims cannot be reconciled with Socrates' principle that under no circumstances must a man do what is unjust or treat people badly (*Crito* 49b-c). However, Woozley does not seem to take into account Allen's point that some laws may be unjust, and that Socrates is not obligated to abide by an unjust law. The very fact that Socrates will disobey unjust laws, like the demands of the Thirty Tyrants, strengthens Socrates' commitment to God, justice, and, when justice happens to align with the law, the state. Socrates' anecdotes and claims allow him to be loyal to God, justice, and the state without provoking discrepancies between the *Crito* and the *Apology* or between Socrates' assertions.

However, Socrates' statement that he would refuse to stop phi-

losophizing if the court were to command it draws our attention back to Socrates' perceived defiance of a direct order from the courts. Socrates' claim in the *Apology* that he would continue philosophizing even if the court ordered against it can be interpreted as either a bold declaration against obedience or a subtle response to a hypothetical situation. If we conclude that Socrates staunchly opposes obedience to the courts' decisions, we encounter a contradiction with the claim in the *Crito* that we must obey all laws. If we view Socrates' assertion as a retort against a hypothetical scenario, the problem is trivialized and has no real standing in the consistency argument since no real action took place on the part of Socrates or the courts. Regardless, Socrates makes his intentions clear on the possibility of the court releasing him under the condition that he no longer practice philosophy:

> If, as I say, you were to dismiss me on that condition I would reply that I hold you in friendship and regard, Gentlemen of Athens, but I shall obey God rather than you, and while I have breath and am able I shall not cease to pursue wisdom or exhort you, charging any of you I happen to meet in the accustomed manner. (*Apology* 29d)

Woozley does not view Socrates' claim of disobedience as an inconsistency. He maintains that because the court order is hypothetical in this instance, Socrates has done no harm by threatening to disobey it. Woozley identifies a difference between the court ordering Socrates to do something and the court telling him what would happen if he actually does it (Woozley 44). Thomas Brickhouse and Nicholas Smith, in *Plato and the Trial of Socrates*, point out that while Socrates was in the military he must have been unable to philosophize. By going to war, Socrates was putting his life, and therefore his ability to philosophize, at risk. Thus, to Socrates, some things must actually be more important than philosophizing, despite his being ordered by God to practice philosophy. By

going to prison and subsequently being executed, Socrates' ability to philosophize will be permanently destroyed. This reasoning, according to Brickhouse and Smith, shows that Socrates would not necessarily disobey the court's order to cease philosophizing in the manner he says he would, because it would result in his death and inability to practice philosophy (Brickhouse and Smith 227). In this way, Brickhouse and Smith reconcile Socrates' threat to disobey a direct order from the courts with his loyalty to justice.

Santas, however, offers an advantage to Socrates disobeying a direct court order to stop philosophizing. Santas asserts Socrates would have reasoned that practicing philosophy, even under strict orders from the court not to, would have benefits. If Socrates were willing to accept the penalty of doing so, he would have justified his actions based on his view that philosophizing is good for the city and that God ordered him to philosophize (Santas 50). In the journal article Socrates and the State, Richard Kraut refutes Santas' assertion and aligns his argument with that of Brickhouse and Smith: By practicing philosophy, Socrates would still be breaking the law. He would subsequently be caught and put to death. The assertions of Brickhouse, Smith, and Kraut appear to hold more weight than those of Santas, given that they rest on one simple fact. By philosophizing while under orders not to, Socrates would be killed and could no longer practice philosophy. Regardless of the benefits of actually practicing philosophy, Socrates' efforts would be short-lived and his subsequent execution would terminate any future hopes of philosophizing. Therefore we can conclude that Socrates' assertion was probably a response to the hypothetical situation. Given the arguments of Brickhouse, Smith, and Kraut, we can also presume that Socrates may not have carried out his threat anyway if the court were to actually order him to stop philosophizing. Having made efforts to resolve Socrates' hypothetical disobedience against the court, it is now imperative to consider what many critics acknowledge as the most formidable adversary

to consistency between the *Crito* and the *Apology*: the personified speech of the Laws.

The speech of the Laws in the latter portion of the *Crito* presents a challenge to those who posit consistency between the *Crito* and the *Apology* and has been dealt with by varying means. The personified Laws appear to contradict Socrates' anecdotes of disobedience in the *Apology*. In the *Apology*, it was determined that Socrates believes that it is acceptable to disobey unjust laws under specific circumstances. Conversely, the personified Laws argue that a state cannot subsist if individuals decide to disobey the law (*Crito* 50b). The Laws also insist that Socrates owes a debt to the state, since the state plays a critical role in his existence:

> In the first place did we not bring you into existence? Your father married your mother by our aid and begat you. Say whether you have any objection to urge against those of us who regulate marriage? [...] Or against those of us who regulate the system of nurture and education of children in which you were trained? (*Crito* 50d-e)

The most elaborate explanation for the consistency of the speech of the Laws with the *Apology* comes from Weiss, who proclaims that the personified Laws cannot represent Socrates' actual views for several reasons: the speech of the Laws was invented for Crito's benefit, Socrates will always fight for the just, and the fact that Socrates credits the speech to an orator. These arguments function as a refutation of the Endorsement Assumption, the presumption most critics make that the Laws represent Socrates' views. Woozley, for example, explicitly deems the view held by Weiss implausible (Woozley 29). First, Weiss argues that Socrates creates the speech of the Laws for the benefit of his friend, Crito, as a last attempt to save his soul. Crito, a wealthy Athenian citizen, offers to fund Socrates' escape through bribery (*Crito* 44c-d). Weiss says that Crito's argument for Socrates' escape is not philosoph-

ical; Crito concerns himself with money and his reputation (considered the concerns of "the many" by Socrates). Socrates points out that the priorities of "the many," namely wealth and fame, are unimportant in comparison to justice and the soul (*Crito* 47c-d). From these passages, Weiss argues that Socrates invented the speech of the Laws to convince Crito that he should be more concerned with justice and the soul than with money and notoriety. Weiss concludes that Crito is "a friend to Socrates' body, but not to Socrates' soul." (Weiss 56) Allen points out along the same lines that the *Crito* claims that the soul can be harmed only for *committing* an injustice, not for *suffering* an injustice (Allen 109). By refusing Crito's offer to escape prison, Socrates has assured that his soul will avoid all damage, irrespective of whether his body will be hurt. Refusing to accept the laws would be committing an injustice, as this was the verdict put forth by the jury. Socrates transcends the situation by concerning himself with the state of his soul, rather than the state of his body (a concern of Crito's). Weiss believes that Socrates is trying to convince Crito to share his same concerns. Weiss also outlines a strong argument for why Socrates will always fight for the just. Referring back to the Leon of Salamis case, Socrates resists an unjust command, despite its legality. Socrates aligns himself with what is just, not what is legal. Weiss brings up Socrates' anecdote in the *Apology* of how he cast his vote against the indictment of the ten generals (32b). In this situation, the case was both illegal and unjust. Socrates' vote in this situation aligned with both the legal and the just. Weiss contends that Socrates will always affiliate with justice, but will also promote the law if it happens to coincide with what is just. Lastly, Weiss points out that Socrates identifies the speech of the Laws as what an orator would say, providing us with a body for which the Laws are speaking. According to Weiss, we can presume that this speech does not reflect the views of Socrates himself, as Socrates says of orators in the *Apology*, "They have said little or nothing

true." (17b)

Masha Marchevsky, in her article "Socrates Misinterpreted and Misapplied: An Analysis of the Constructed Contradiction between the Apology and the Crito," addresses two problems with Weiss's interpretation of the speech of the Laws. First, Marchevsky posits that Weiss's characterization of Crito as unintelligent is unwarranted. She defends Crito's pervading confusion in the dialogue as a lack of understanding regarding how Socrates reasons. Secondly, Marchevsky disagrees with the idea that Socrates would lie about his beliefs for Crito's benefit. Woozley, however, while adhering to the Endorsement Assumption, takes a different approach to attempt to salvage consistency between the *Crito* and the *Apology*. Woozley indicates that, in the *Apology*, Socrates professes individuals have the right (or even duty) to disobey the law. According to Woozley, this is in stark contrast to the *Crito*, which allows, based on the Laws, only the limited freedom of persuading the authorities of the injustice of a particular law. Known as the "persuade-or-obey" doctrine, Woozley, with later commentary from Kraut, attempts to use this argument to reconcile the apparent contradiction between the *Apology* and the *Crito* brought on by the speech of the Laws.

In the *Crito*, the personified Laws famously state, "You must do whatever your city and country orders, or you must persuade it in accordance with where justice really is." (51b) This statement opens up debate for whether an attempt at persuasion can be considered an alternative to actually obeying the law. Rights activists such as Martin Luther King, Junior have used this idea as a justification for civil disobedience. Woozley places an emphasis on the "or" in the Laws' assertion. He argues that, based on the Laws' speech, a man must persuade or obey, not obey until he can persuade. Woozley also points out that persuasion might not necessarily take the form of blatant disobedience, but can serve as an introduction to the idea of civil disobedience. Woozley uses this

concept of civil disobedience to justify Socrates' challenge to the court in the *Apology*, on the hypothetical issue that Socrates be released under the condition that he stop philosophizing (Woozley 30). Kraut, on the other hand, opposes Woozley and professes a more liberal view of the doctrine. He believes that Woozley's idea of civil disobedience is anachronistic, as there were no policies in Athenian law that allowed for persuasion to change the laws. To replace Woozley's argument, Kraut proposes that citizens need only make a credible attempt to persuade the appropriate authorities to change the law (Kraut 661). While Kraut successfully refutes Woozley, his own explanation is also problematic. The word "persuade" is generally taken to be a "success verb," which refers not only to an action, but also to the successful outcome of that action. Simply making a legitimate effort to persuade the authorities is not enough to justify disobedience. Thus it would appear that the persuade-or-obey doctrine does not hold as interpreted by either Woozley or Kraut and can also not be used as a means to justify civil disobedience.

By dissecting various intricacies of different authors' arguments, we can generate a coherent view of why the *Apology* is indeed consistent with the *Crito*. While Marchevsky brings into question much of Weiss's argument regarding the true purpose of the personified speech of the Laws, Weiss still makes the legitimate point that Socrates always aligns himself with what is just. Allen agrees that Socrates' priorities focus not on preserving his body, but on preserving his soul. Socrates accomplishes this by consistently positioning himself with justice, regardless of whether justice and legality line up. Allen also points out that Socrates will disobey an unjust command if it does not coincide with what is just, as was the case with Leon of Salamis. With the proposal that Socrates is loyal to justice, we discover that there is no incongruity between the *Crito* and the *Apology*.

George Grote's view is that the *Crito* and the *Apology* are dia-

metrically opposed. His claim is based on the idea that after writing the *Apology*, Plato realized that his portrayal of Socrates in the dialogue was akin to that of a lawless criminal. According to Grote, Plato wrote the *Crito* to rectify these perceptions. This view is hard to accept because Plato would likely have been aware that Socrates made efforts to be meticulously consistent to his own principles. Plato would not have violated such a basic principle of Socrates' to try to rectify an apparent misconception. Socrates asserts in Plato's *Gorgias* (428c): "As for myself, I would rather that my lyre were out of tune, or a choir I was training, and that the greater part of mankind should dissent from me and contradict me, than that I should be out of tune with my own single self." This statement provides us with additional proof that, based on Socrates' own principles and our faith in Plato's writing, Socrates' lyre is not out of tune and he is in fact consistent with himself.

References

[1] Allen, R. E. *Socrates and Legal Obligation*. Minneapolis: University of Minnesota, 1980.

[2] Brickhouse, Thomas C., and Nicholas D. Smith. *Plato and the Trial of Socrates*. New York: Routledge, 2004.

[3] Grote, George. *Plato and the Other Companions of Sokrates*. Bristol: Thoemmes, 1992.

[4] Kraut, Richard. "Review: Plato's Apology and Crito: Two Recent Studies." *Ethics* 91.4 (1981): 651-64.

[5] Kraut, Richard. *Socrates and the State*. Princeton: Princeton University, 1987.

[6] Marchevsky, Masha. "Socrates Misinterpreted and Misapplied: An Analysis of the Constructed Contradiction between the Apology and the Crito." *Macalester Journal of Philosophy* 13.1 (2004): Article 4.

[7] Panagiotou, Spiro. *Justice, Law and Method in Plato and Aristotle*. Edmonton: Academic Printing, 1987.

[8] Plato. *Gorgias*. Trans. Terence Irwin. Oxford: Clarendon, 1979.
[9] Santas, Gerasimos. *Socrates: Philosophy in Plato's Early Dialogues*. London: Routledge & Kegan Paul, 1979.
[10] "The Internet Classics Archive | Apology by Plato." The Internet Classics Archive: 441 Searchable Works of Classical Literature. Web. 25 Mar. 2011. <http://classics.mit.edu/Apology/crito.html>.
[11] "The Internet Classics Archive | Crito by Plato." The Internet Classics Archive: 441 Searchable Works of Classical Literature. Web. 25 Mar. 2011. <http://classics.mit.edu/Plato/crito.html>.
[12] Weiss, Roslyn. *Socrates Dissatisfied: an Analysis of Plato's Crito*. New York: Oxford University Press, 1998.
[13] Woozley, A. D. *Law and Obedience: the Arguments of Plato's Crito*. Chapel Hill: University of North Carolina, 1979.

THE RESOLUTION OF INCOMPATIBILITIES IN KANT'S COMPATIBILISM

Jared Clifton

Texas A&M University, 2012

Compatibilism is a position which arises in view of the dilemma commonly referred to as the problem of free will. This problem is predicated on the conflicts which arise as a result of certain representations of free will and determinism, both of which are variously construed. The concepts at play in these discussions implicate some of our most fundamental intuitions concerning the choices we make, and the world we inhabit. In Kant, there is a great sensitivity to these considerations, and clear indication that he too was profoundly concerned with the antagonism which seems to occur. In this exposition, it will be my endeavor to defend the position that Kant was a proponent both of determinism and the freedom of human will, and thus ought to be considered, by contemporary standards, a compatibilist.[5] Much must be done in the way of explication before presenting my reasons for interpreting of Kant such a thing. Foremost, it will be helpful to set myself straight with respect to the angle with which I will approach the terms 'determinism' and 'free will'.

The free will problem, as I understand it, can be derived in several forms. I've selected the one which I feel is most clear and concise; still, it must be granted that any rendition of this problem requires some static metaphysical notions. Suffice it to say, there is a vast literature of mutually exclusive opinions on the subject of free will: of its definition, of its existence, and of the abundant topics upon which it comes to bear; but for the sake of argumentation (and brevity), I will more or less blindly affirm that in order for any free will thesis to be respectable, it must meet the following two stipulations: (i) free will entails that agents have the ability to act otherwise than they end up acting; and (ii) free will is a sufficient condition for moral responsibility. In addition to these

5. Kant's views with respect to these subjects vary at times considerably between his early work and his later work; to mitigate this, I limit the scope of my investigation to *Critique of Pure Reason*, *Idea for a Universal History with a Cosmopolitan Purpose*, and *Groundwork of the Metaphysics of Morals*.

distinctions, it would be good to convey exactly what I mean by causal determinism; in short, this will refer to the ontological position that all events are dictated by previous events and/or the natural laws. With these indulgences in mind, consider the following set of propositions:

(I) Human agents have free will.
(II) Causal determinism obtains.
(III) Human agents have free will only when causal determinism does not obtain.

There are a number of ostensibly reasonable positions available with respect to how one might take up or discard these propositions, but it seems clear enough that holding all three simultaneously gives rise to contradiction. As it is, assuming that contradictions are an unwanted phenomenon, there are three fashionable stances taken in the contemporary debate: libertarianism, which claims that we have superior evidence to maintain the truth of (I) over (II), but is immediately faced with the demand of making intelligible some metaphysical position that would falsify (II)[6]; hard determinism, which claims that we have superior evidence to maintain the truth of (II) over (I), but is immediately faced with the demand of either doing away with or rehashing quite a few of our intuitions that seem to depend on the truth of (I)[7]; and compatibilism, which denies the need to reject (I) and (II), but must face the demand of providing good reasons for rejecting (III). In order to properly defend my thesis, I must successfully illustrate the truth of three distinct claims about Kant's thinking: that there

6. I'm thinking of something like agent causation, which is a view that Kant discusses in the *Critique of Pure Reason*; see Aristotle and Epicurus for agent-causal libertarianism defended; the position, to be scandalously brief, is orbits the proposition that agents can initiate previously undetermined causal chains.
7. I'm referring to intuitions about: moral responsibility, blame, praise, etc.

is good reason to believe that he espouses causal determinism; that there is good reason to believe that he considers the human will free; and that there is good reason to think that he views the proposition that they are incompatible false. Having presented a method of distinguishing compatibilism from the other prominent views, I can now go about elucidating why it is that I think Kant ought to be thought a member of the former.

Throughout Kant's work, it is apparent that he was tremendously concerned with the nature of human knowledge; one prevailing interest in his work is predicated on the question: "What kinds of things, if any, can the human mind know, and how?" For Kant, as he argues in the *Critique of Pure Reason*, our knowledge is very sternly restricted to facts of a conceptual or mathematical nature, and empirically demonstrable facts about the natural world. One significant consequence of this position is the that the human mind is unable to draw inferences involving what he calls the "speculative method", or those applications of reason which infer metaphysical propositions. The justification he provides for this assertion pertain to the dynamic and active role he determines the mind must play in coming to know anything at all. Kant, I submit, in contemplating the notion of a *subject* of experience which is presented with a universe which is an *object* of experience, derived the conclusion that in such a circumstance, the subject must have perceptional, intellectual, and conceptual capabilities of some kind. Notice the important implication of this proposition – one which I will try not to butcher – human minds are not passive with respect to their coming to know the features of their experience. This, Kant explains, limits applications of reason to only those facts which are within the empirical realm of space and time. Kant, though, was not claiming that other kinds of knowledge could not come about (this would be highly compromising for the enterprise of philosophy!). Given that humans are perceiving things, he argues, and that we can gather from this

that there are certain limits that are imposed on our ventures in perceiving, he posited that if the mind deliberates long and hard enough, it can decipher truths of a kind which he terms the "synthetic a priori". This form of knowledge stands apart from synthetic a posteriori knowledge, which is composed of the matters of fact we come to know through experience, and analytic a priori knowledge, which is composed of necessary logical truths and of unambiguous matters of definition.[8] I think I'm guilty of jumping the gun a bit, so allow me to backtrack and clarify what I understand Kant to say in distinguishing between synthetic and analytic: judgments that are synthetic are those whose predicates are not contained in their subjects; and, contrariwise, analytic judgments are those whose predicates are contained in their subjects. These synthetic a priori types of knowledge, Kant goes on to say, exist in two categories: those which compose the "form of experience", and those which compose the "form of understanding". A "form of understanding" involves the utilization of the conditions of possible understanding in order to show that certain fundamental principles must obtain in the world of appearances; one such principle, important to Kant's thesis as well as mine, is the principle of causal determinacy. He writes:

> The principle of an unbroken connection between all events in the phenomenal world, in accordance with the unchangeable laws of nature, is a well-established principle of transcendental analytic which admits of no exception.

8. Analytic a posteriori inferences don't really happen, since there isn't any need to appeal to experience in these circumstances

"Forms of experience", in a parallel manner, are utilized by Kant to establish the notions of time and space[9], which are *only precisely* that to which human experience must conform – that is, time and space cannot in principle be said to apply to "das ding an sich" (the thing in itself), but rather, only to the proper objects of knowledge, which he called "phenomenon". It is simply a matter of fact, he claims, that humans exist in a state of constructing all of their experiences with the qualities of temporal and spatial extension. Herein the prevention of metaphysical speculation again rears its head, since the possibility of making inferences about facts which have nothing to do with our spatial-temporal experiences, given our constraints as perceiving things of the specific kind that we are, is absurd. While it seems that the case has been made with some sufficiency that Kant holds to the truth of (II), the following troubling question arises upon reflection of the methods employed: Does he then[10] arrive at a notion of morality by similar means – that is, by pure reason? Or perhaps even more troubling: If morality is not the kind of thing that has spatial-temporal extension, is it possible on Kant's account to make intelligible such a thing as knowledge of morality? In the following paragraphs, I will try to convey what I see Kant providing as answers to these queries.

It's important to note, firstly, that for some of Kant's predecessor's[11], the project of establishing our beliefs about morality and theology required providing a system of metaphysics which allow for such a thing. Kant, on the other hand, restricting the realm of

9. For Kant, arithmetic and geometry can also be inferred in this way.
10. This representation of the flow of ideas is very artificial, since Kant in no way developed his ideas in this manner. I'm merely trying to express the sense in which, since Kant's thought is so intricately webbed together, it's almost as if one must arbitrarily designate a starting point and precede form there in order to penetrate it.
11. This is a rather large claim, and one which I probably shouldn't have included, but from my experience of the empiricists prior to Kant, this seems to be the case; one notable example being Berkeley.

metaphysics to a thing which is unknowable (and improvable) by our limited perceiving existence, flips this procedure on its head, in a sense. He instead endorses the position that we are not only entitled, as limited perceivers, to beliefs with respect to the moral and theological categories, but that we are in some sense unavoidably aware of the truth of certain moral and theological propositions. These convictions, then, lead inexorably to certain metaphysical propositions, which can be examined by pure reason. Thus, as I see it, the process of coming to justify our beliefs – and I use this term loosely here, because as I've stated before, Kant thinks they lie outside the scope of pure reason – about such things is not a question of rational justification, but of primitive belief, in much the same way that, in philosophy of logic, one must come to accept as least some of the intuitive propositions concerning our capacity to reason. As Kant so eloquently put it:

> Happiness, therefore, in exact proportion with the morality of rational beings (whereby they are made worthy of happiness), constitutes alone the supreme good of a world into which we absolutely must transport ourselves according to the commands of pure but practical reason.

Practical reason, for Kant, is nigh inseparable from the notion of human will, and is primarily applied with respect to the questions of morality, or how we out to act. I am trying to be careful, though, not to suggest that the matter of morality is divorced altogether from pure reason; I am only claiming that, for Kant, the basic facts of moral experience are readily available to all persons. One of the immediately necessary consequences[12] of this fact, Kant claims, is the freedom of human will, as evidenced by his statement in the *Critique of Practical Reason* that "freedom actually exists, for this idea is revealed by the moral law"; for if morality is to have

12. Or perhaps it would be better to say "necessary conditions".

any cogency whatsoever, it must be the case that at any moment, we could in fact have done otherwise. And in another place he notes, along a similar line of this reasoning, that a moral agent "judges that he can do a certain thing because he is conscious that he ought, and he recognizes that he is free, a fact which, but for the moral law, he would never have known." It seems now that I have provided an at least very basic understanding of how Kant arrives at justification for (I) and (II), but it remains unclear how he intends to provide good reasons to reject (III). Kant, in fact, at numerous locations throughout his works, points out the apparent antagonism which arises upon espousing these two positions, and the manner in which his system, which does espouse them, might go about dissolving said antagonism.

If the faculty of pure reason is applied to the issues between free will and causal determinism, the discovery is straightaway made by Kant that both the proposition of free will and it's negation can be provided with compelling proofs. In his subsections "on the thesis" and "on the antithesis" in his section "On the Third Antimony", he goes about providing an account of this. Therefore, pure reason cannot assist in the mitigation of the problem of free will. How then to approach it? I find that the critical issue lies in a distinction which I've above only superficially rendered, and that is the distinction Kant makes between the world of appearances, and of things-in-themselves. If the proper residence of empirical, scientific inquiry is only in the world of appearances, as is required by the nature of our faculties, there seems to be room, for Kant, to speculate about that which is not necessary, but is possible; not only possible, in this case, but confirmable by experience. Notice that according to Kant's account, we have no good theoretical justification – that is, a series inferences arrived at by pure reason – for concluding (or rejecting) the proposition that causal determinism is true of things in themselves. This is due to the fact that causal determinism is a position arrived at, as stated

above, by the forms of the understanding that humans come to have about the world of appearances. On the other hand, since free will is a necessary (and sufficient) condition for morality, it seems that we have good practical reasons to take as a matter of faith that the human will is indeed free. In a similar way, parallel to this line of reasoning, Kant holds the belief that in order to speculate about any of those facts which are not proved or denied by pure reason, e.g. divine beings, the soul, immortality, they must occur as a result of our fundamental moral experiences, and the elaboration of those primitive beliefs at the hands of pure reasoning. Hence, if my interpretation is correct, Kant seems to provide compelling reasons for rejecting (III). It would be good, in view of this rendition of Kant's thought, to test all of the aforementioned interpretations against one of Kant's major historical projects, *Idea for a Universal History with a Cosmopolitan Purpose*, in which he employs many of these concepts.

In the opening sections of *Idea for a Universal History with a Cosmopolitan Purpose*, Kant puts his cards on the table when he notes that:

> whatever concept one may hold, from a metaphysical point of view, concerning the freedom of the will, certainly its appearances, which are human actions, like every other natural event are determined by universal laws.

Kant then sets out his purpose to be the determination by use of pure practical reasoning that from the "regular movement" of human history, some fact about its "slow evolution" might be derived. Further, Kant continues, the amalgamation of our various knowledge of the world of appearances and of those beliefs arrived at about certain things-in-themselves by practical reason allows us to speculate on the ends human-kind has in store for itself at the hands of Nature. This thing, Nature, has given rise not only to the "natural capacities" of "initial" humans, in individuals and

in how they will come together to form societies, which allow for their evolution – and ascension – but also the well-ordered and causal qualities of the world of appearances which humans experience. But if Nature acts as an ordering force upon the world of appearances, it seems as if this thing composes a part of the whole of the world of things-in-themselves. How then can we come to know anything about it, such that we can go about employing it to the purpose which Kant has outlined? Herein, I think, there is reason for assurance that my rendition of Kant's position may be correct; for in his Ninth Thesis he states that:

> A philosophical attempt to work out a universal history according to a natural plan directed to achieving the civic union of the human race must be regarded as possible and, indeed, as contributing to the end of Nature.

To put it explicitly, the idea of that we can come to a notion of human history where this Nature obtains (because of the allowances Kant's system makes with respect to the possibilities of experience and the applicability of practical reasoning) is useful not only because it serves to "[clarify] the confused play of things human," but also for:

> giving a consoling view of the future (which could not be reasonably hoped for without the presupposition of a natural plan) in which there will be exhibited in the distance how the human race finally achieves the condition in which all the seeds planed in it by Nature can fully develop and in which the destiny of the race can be fulfilled here on earth.

This selection, though lengthy, is worth its space due to the degree to which it exemplifies my above construal of Kant's reasoning. Nature, in the sense just qualified, allows our coming to comprehend a guiding thread in history. We saw in the *Critique of Pure Reason* a similar principle – namely, that "They ['transcendental

ideas, which are of some modi of the pure conceptions of reason'] follow the guiding thread of the categories," which, for Kant, are the most general concepts, in terms of which every object must be viewed in order for it to become an object of empirically derived knowledge. So it seems, in much the same way that our coming to have moral convictions through practical reason resists the possibility of moral relativism (or worse, moral nihilism), our convictions with regard to the movement of history which we also arrive at through practical reasoning help resist the "[destruction of] all practical principles, and [the making of man by] Nature...[into] a contemptible plaything".

In summation, it seems that there is sufficient reason to believe that Kant espouses causal determinism, that he considers the human will free, and that he views the proposition that they are incompatible false. In view of this, I feel I have adequately shown that Kant does seem to, in view of my interpretation of his thought, espouse the position which we would today call compatibilism.

References

[1] Kant, Immanuel. "The Critique of Practical Reason."

[2] Kant, Immanuel. "Critique of Pure Reason."

[3] Kant, Immanuel. "Groundwork for the Metaphysics of Morals."

[4] Kant, Immanuel. "Idea for a Universal History from a Cosmopolitan Point of View." On History. 1st Edition. Lewis White Beck. New York: Macmillan Publishing Company, 1963. 11-26. Print.

SUICIDE AS A RESPONSE TO THE REALIZATION OF THE ABSURD: A MISTAKE IN CAMUS' REASONING

Sarah Baver

St. Lawrence University, 2012

In this essay, I will re-construct Albert Camus' philosophical inquiry into the absurd in the chapter "An Absurd Reasoning" in his essay *The Myth of Sisyphus* by discussing how his reasoning that suicide is not the proper response to the absurd collapses in on itself, and by demonstrating how he does not fully explicate the distinction between life being worthless and life being meaningless that he intimates. In this chapter of his essay, Albert Camus contemplates the notion of the absurd. His discussion of the absurd centers on the question, "does the Absurd dictate death?"[13] This question asks whether or not the realization of the absurd can cause a person to commit suicide. Camus comes to the conclusion that suicide is not the proper response when a person is confronted by the absurd since suicide conforms to the absurd rather than revolts against it.[14] Following Camus' reasoning, the realization of the absurd is inaugurated by an, "impulse of consciousness."[15] This impulse of consciousness not only incites a confrontation with the absurd, but also maintains that confrontation. In turn, this conscious awareness then both inflames and sustains a personal revolt of passion against the absurd.[16] Committing suicide because of the absurd, then, is not the proper response since realization of the absurd should instead provoke a passionate revolt against it.

But, Camus' argument that suicide is not the proper response to the absurd is mistaken. According to Camus, suicide is, "rarely committed through reflection."[17] The logical consequence of this statement, then, is that suicide cannot be incited by a confrontation with the absurd since such a confrontation requires conscious awareness, and that which prompts suicide lacks this awareness.

13. Albert Camus, *The Myth of Sisyphus*, "An Absurd Reasoning; (NY: Vintage Books, 1955), 7.
14. Ibid., 40.
15. Ibid., 10.
16. Ibid., 41.
17. Ibid., 4.

Since suicide is not committed through reflection, it cannot, therefore, be caused by a realization of the absurd. Thus, suicide cannot be a response to the absurd, let alone a proper one.

With this in mind, I will argue in this essay that suicide is never a response to the realization of the absurd, but rather is a response to the vague feeling of absurdity. This vague feeling of absurdity is that which provokes the belief that life is worthless. The belief that life is worthless, and not the belief that life is meaningless, can cause a person to commit suicide. In order to support my argument, I will draw upon Keiji Nishitani's *Religion and Nothingness* to further demonstrate how the realization of the absurd can serve only as a catalyst for a person to passionately revolt against the absurd, and never as a catalyst for committing suicide. With this mistake in Camus' argument established, I will use Simone de Beauvoir's *The Ethics of Ambiguity* to further elucidate how a person's lack of reflection when confronted by the vague feeling of absurdity and compounded with the frustration begotten by the ambiguity of his consciousness, can lead him to deem his life worthless, and can thus incite him to commit suicide.

In the beginning of "An Absurd Reasoning" Camus states the fundamental aim of all philosophical inquiry: "Judging whether life is or is not worth living."[18] This is the framework that structures Camus' discourse on the absurd. The focal point of this discourse is to answer a question which is at the heart of determining whether life is or is not worth living, "does [life's] absurdity require one to escape it through hope or suicide?... Does the absurd dictate death?"[19] This question is prompted by mistakes in reasoning, such as, "people have... pretended to believe that refusing to grant a meaning to life necessarily leads to declaring that it is not worth living."[20] With the acknowledgment of the inherent

18. Ibid., 3.
19. Ibid., 7.
20. Ibid.

flaw in this thinking, Camus brings to light the distinction between life being absurd, or meaningless, and life being worthless. The main concern of his essay, then, is to fully illuminate all that encompasses the realization of the absurdity of life. This involves not only identifying the spark which awakens this realization, but also determining the value in choosing one of the two possible implications of that realization creates: "suicide or recovery."[21]

To begin this inquiry into the absurd, Camus describes the feeling of absurdity as that which grips an individual when he finds himself, ". . . in a universe suddenly divested of illusions and lights" and "feels an alien, a stranger."[22] It is this, "divorce between man and his life, the actor and his setting,"[23] that causes the feeling of absurdity to permeate an individual's life. When a person feels the absurd, he feels anxious.[24] But, this feeling of anxiety, "the feeling of the absurd is not. . . the notion of the absurd."[25] When man is, "lost in the world and its diversions, anxiety is a brief fleeting fear."[26] That feeling of anxiety, though, transforms from a fleeting fear into a sustained experience when it, "becomes conscious of itself."[27] By being conscious of itself, that feeling of anxiety morphs into the experience of anguish, and it is that experience which causes a person to confront the notion of the absurd.

Thus, when a person is conscious of his anxiety, he fully realizes the absurdity of life. This realization can be described as the understanding that, "the absurd is essentially a divorce."[28] More specifically, this divorce is one between the world and man – the world and man may be separate entities, but when man is within

21. Ibid., 10.
22. Ibid., 4.
23. Ibid.
24. Ibid., 18.
25. Ibid., 21.
26. Ibid., 18.
27. Ibid.
28. Ibid., 23.

the world it is the absurd which links him to it. In this way, the absurd necessarily arises out of the confrontation between man and the world. As Camus writes, "I can therefore say that the absurd is not in man nor in the world, but in their presence together."[29] Camus contrasts this with the example of being "a tree among trees" or "a cat among animals"[30] – if one were to exist as such, then he would belong to the world. But, as Camus points out, this is not the case with man; he does not belong to the world, but rather to himself. Despite this, though, man exists within the confines of the world, and because of that relationship, the absurdity of life is a given for man. When man does become consciously aware of his place within the world, and therefore also of the given of the absurdity of his life, he is thus "forever bound" to the absurd.[31] This bounding to the absurd can be further explained by examining that which gives rise to it: conscious awareness.

Camus devotes much of "An Absurd Reasoning" to discussing how this conscious awareness of the absurd is attained. This awakening often takes place when the weariness and repetitiveness of a mechanical life becomes overwhelming. More specifically, when such a life becomes overwhelming it is when a person is consumed by his anxiety so much that it is transforms into anguish. Camus provides his reader with an excellent example of such a life: a person spends every day, Monday through Saturday, performing the same actions. He always wakes up in the same bed, has his meals at their appointed times, travels to and from work by the same means, and does the same tasks when he is at work.[32] The moment a person becomes aware of the extent to which his life is mechanical is when he suddenly questions himself – he asks himself, "why?"[33] This self-questioning not only marks the

29. Ibid.
30. Ibid., 38.
31. Ibid., 24.
32. Ibid., 10.
33. Ibid.

ending of the weariness of the acts of a mechanical life, but also marks a beginning which, "inaugurates the impulse of consciousness"[34]. Camus describes this inauguration of consciousness as an "awakening"[35], and the consequence of this awakening can be one of two possibilities: suicide or recovery.[36] Before explaining what these two consequences entail, though, we can turn to the writings of Keiji Nishitani to further elucidate how the act of questioning oneself, the asking of "why", enables a person to end the weariness of his mechanical life.

In *Religion and Nothingness*, Nishitani discusses the religious quest. Nishitani writes that, "it is a mistake to ask 'What is the purpose of religion for us' " because that is a question which obscures a person from questioning himself.[37] Questioning oneself is central to the religious quest, and attempting to understand religion apart from this question, "betrays an attitude,"[38] of trying understand religion apart from the quest it unlocks for the individual. Nishitani writes that, "there is no other road that can lead to an understanding of what religion is and what purpose it serves," than asking the breakthrough question, "for what purpose do I myself exist?"[39] Once a person asks himself this question, he can undertake the religious quest.

Similar to Camus, Nishitani discusses how a person comes to question himself. As Nishitani has shown, questioning oneself and embarking on the religious quest are the one in the same. Because of this, a person is able to question himself when the question of, "what purpose does religion serve?" ceases to have meaning. This question ceases to have meaning when it no longer needs

34. Ibid.
35. Ibid.
36. Ibid.
37. Keiji Nishitani, Religion and Nothingness, "What is Religion" (CA: University of California Press, Inc.), 340.
38. Ibid.
39. Ibid.

to be asked, and this occurs once religion becomes, "a must for life."[40] Religion becomes essential when, "everything else [in life] loses its necessity and utility."[41] These other things are the engagements we create to give some kind of semblance of meaning and purpose in our lives. As with Camus, these engagements can give rise to living a mechanical life. And, as with Camus, Nishitani points out, "meaninglessness... lies in wait at the bottom of those very engagements that bring meaning to life."[42] A person realizes this when weariness seeps in from performing those same engagements continually. Nishitani identifies this moment as, "the point at which that sense of nihility, that sense that 'everything is the same'... brings the restless, forward-advancing pace of life to a halt and makes it step back."[43] This sense of nihility is analogous to the transition Camus distinguishes as shifting from feeling anxiety to experiencing anguish. As Camus has shown, when a person becomes overwhelmed by the weariness of his mechanical life, he reaches a breaking point. This breaking point is marked by anguish, and this experience of anguish opens up the possibility of confronting the notion of the absurd. Both Camus and Nishitani explain how the appearance of anguish and nihility signal, "that one's awareness of self-existence has penetrated to an extraordinary depth."[44] The consequence of this self-awareness for both Camus and Nishitani is questioning oneself, and this questioning allows a person to put an end to the weariness of his mechanical life. For Nishitani, this questioning also allows a person to embark on the religious quest, while for Camus it allows a person to respond to the absurd through either the recovery of revolt or through the act of suicide.

40. Ibid.
41. Ibid.
42. Ibid., 341.
43. Ibid.
44. Ibid.

With the source and importance of this self-questioning established, we can turn to discussing the first of the two consequences it triggers, which Camus identifies as recovery through revolt. Camus comments on the revolt against the absurd this questioning incites when he writes, "... the very thing that lead to despair of the meaning and depth of this life now gives it its truth and its clarity,"[45] and, "that revolt gives life its value."[46] The reason why this revolt provides a person with clarity and value is because after becoming conscious of the absurdity of life, everything in one's life is upset. This upsetting inflames a person to passionately deify the absurd every day. As Camus has shown, once a person becomes aware of the absurd, he is forever bound to it. Just as a person cannot escape the absurd because it necessarily arises out of his placement within the world, so too can he not escape the absurd once he becomes consciously aware of it. Thus, when a person does confront the absurd, he will remain conscious of it. And this conscious awareness is continually acted upon through the defiance of the absurd, that is, through the revolt of living in spite of the absurd. In turn, this revolt sustains a person's conscious awareness of the absurd. This is why a person becomes forever bound to the absurd – his revolt against the absurd sustains his conscious awareness of it, and this conscious awareness thereby sustains his revolt against the absurd. In other words, his actions are simultaneously both a response to and a reaction against the absurd of which he is consciously aware.

It is crucial to note here that this circle of action with conscious awareness of the absurd at the center of it does not trap an individual inside of it. Instead, this conscious awareness of the absurd discloses a person's being. By disclosing one's being, a person is able to own his being. This is where the recovery through the

45. Albert Camus, *The Myth of Sisyphus*, "An Absurd Reasoning" (NY: Vintage Books, 1955), 28.
46. Ibid., 40.

revolt against the absurd lies – the revolt empowers a person to own every action he makes. By owning every action one makes, a person is able to recover from the weariness and anguish which seized him when he was caught in the confines of a mechanical life. In this way, when Camus writes of a recovery from the absurd, he is writing of a recovery from the realization that the aspects which guided his mechanical life were ultimately meaningless. The consequence of this recovery is that it mobilizes a person to disclose his being, his freedom.

Now that the first of the two possible reactions to confrontation with the absurd has been explained, we can turn to discussing the second reaction Camus identifies in "An Absurd Reasoning", which is suicide. For Camus, "dying voluntarily implies that you have recognized, even instinctively, the ridiculous character of that habit, the absence of any profound reason for living, the insane character of that daily agitation, and the uselessness of suffering."[47] This is why he distinguishes suicide as a consequence of the awakening which occurs after becoming aware of one's mechanical life. For Camus, suicide is an admission of realizing the weariness of living a mechanical life, and thus is also a response to the absurd since that realization of weariness causes a person to question himself and thereby confront the absurd.

Although Camus identifies suicide as a consequence of realizing the absurd, he argues that it is not the proper response. Suicide is not the proper response to the absurd for Camus because, "suicide settles the absurd. It engulfs the absurd in the same death."[48] Since the absurd necessarily arises out of the divorce between man and the world, to remove one part of it would negate the absurd. Thus, when a person kills himself, he also destroys the absurd. The purpose of becoming consciously aware of

47. Ibid., 5.
48. Ibid., 40.

the absurd is to revolt against it, and suicide prevents a person doing exactly that. As Camus writes, "... the point is to live."[49]

Camus may argue that suicide is not the proper response to the absurd because it denies a person the ability to revolt against it, but his argument overlooks a crucial point he makes in "An Absurd Reasoning", which is that suicide is rarely committed through reflection.[50] If suicide is committed during a moment of lapsed awareness, then it necessarily cannot be a response to the absurd as Camus argues. Camus makes it clear that there are two possible consequences to the awakening of conscious awareness of the absurd: recovery or suicide. So, in order for a person to respond to the absurd, he must be consciously aware of it. This is not the case with suicide. Because of this, suicide cannot even be a response to the absurd, let alone the proper one. In this way, conscious awareness of the absurd can only incite a person to passionately revolt against it since that consciousness both sustains the realization of the absurd and the continual reaction against it. As Nishitani points out, the self-questioning which triggers the realization of the absurd enables a person to embark on the religious quest. In this way, the self-questioning which Camus identifies has the same outcome – it enables a person to disclose his being through continuously owning his actions as he revolts against the absurd. With this in mind, the only response to the realization of the absurd is recovery through revolt. This conclusion, though, raises the question of, "if suicide is not a possible response to the absurd, then to what is it a response?"

In "An Absurd Reasoning" Camus states that suicide implies a person has recognized, "the absence of any profound reason for living,"[51] even if it is only instinctively. The key to this statement, though, is that this recognition is one which is not brought

49. Ibid., 48.
50. Ibid., 4.
51. Ibid., 5.

to light fully; it is an instinctive, not a conscious, one. Because of this, what Camus is intimating with this statement is the distinction between life being worthless and life being meaningless. At the moment when suicide becomes a solution, a person believes life is worthless. This belief is brought about by a lapse in conscious awareness; if the person were fully aware of the reality of his situation, he would realize that life is meaningless, that it is absurd. The state a person is caught in when he decides to commit suicide is the vague feeling of the absurd and not the notion of it. This feeling of the absurd is what Camus identifies as the feeling of anxiety; a, "brief, fleeting fear."[52] As long as this fear does not become conscious of itself, it can incite a person to commit suicide. This feeling can be a catalyst for suicide precisely because it is brief and fleeting – it can overwhelm a person to the point of believing suicide is a solution to that moment of anxiety. This is why suicide is committed without reflection – it is an action performed when a person treats himself as a means to an end, that is, a solution to his own self. Because of this, suicide causes a person to trap his being instead of disclosing it.

We can turn to Simone de Beauvoir's essay *The Ethics of Ambiguity* to further elucidate how this brief, fleeting fear of anxiety can incite a person to commit suicide. In this essay, Beauvoir discusses why consciousness is ambiguous. She argues that our consciousness is in a constant state of ambiguity because we desire to ascertain the meanings in the world as well as author those meanings. In this way, we are meaning-discovering and meaning-making beings, and our consciousness is ambiguous because we

52. Ibid., 18.

can never fully satisfy either of these desires. When a person embraces this ambiguity, he is, "raised to the height of moral freedom," because he takes his existence, "as an end through the disclosure of a particular content."[53] Beauvoir is careful to make the distinction between ambiguity and absurdity. For Beauvoir, embracing the ambiguity of consciousness is a revolt against the absurd because it is to, "assert that... meaning is never fixed, that it must be constantly won."[54] This is what the continuous revolt against the absurd grants for the individual: to constantly choose his actions and to own his being, to disclose it.

When a person does not embrace this ambiguity and seize upon the revolt which the realization of the absurd inflames, his being is trapped in a state of confusion and anxiety. He is confused because he is not fully aware of the ambiguity of his consciousness; instead, he is overwhelmed and frustrated by the confusion of which the ambiguity of consciousness gives rise. In this way, Beauvoir's discussion of the ambiguity of consciousness provides Camus' distinction of the brief, fleeting fear of anxiety with a source. Thus, when a person is gripped by the belief that suicide is a solution, his being is trapped in a state of anxiety and confusion which causes him to believe his life is so worthless that he would treat his own being as a means to an end, that is, death. Not only does this trapping prevent a person from fully realizing both the ambiguity of his consciousness and the absurdity of his life, it also blocks him from ever disclosing these truths of his being. It blocks him from disclosing these truths because the act of suicide limits a person to being a means to end rather than being an end in itself.

It is clear from Beauvoir's writings that suicide would pose a serious philosophical problem for her. The same can be argued

53. Simone De Beauvoir, *The Ethics of Ambiguity*, "Ambiguity and Freedom", (NY: The Philosophical Library, 1948), 32.
54. Ibid., 129.

from Nishitani's writings, and it is explicitly evident in Camus' writings that suicide is one of the most significant philosophical problems. When these three writers' discussions of existence and being are considered, it is apparent that suicide is an action that treats a person as a means to an end, and therefore is one that traps a person's being by denying him the possibility to disclose it through the realization of the ambiguity of his consciousness and the absurdity of his life. This understanding, when coupled with Camus' discussion of the absurd demonstrates how his argument that suicide is not the proper response to the realization of the absurd is mistaken. Although Camus' discussion of the consequence of recovery through revolt against the absurd is a thorough one, he overlooks a crucial philosophical truth in his discussion of how suicide is not the proper response to the absurd once the absurd is realized. That truth is that suicide is committed in a moment of failure to be consciously aware, and because of that truth, suicide cannot be incited by the realization of the absurd, and thus cannot be a response to the absurd. Beauvoir's essay *The Ethics of Ambiguity* helps to elucidate why a person would commit suicide if suicide is not incited by the realization of the absurd, and is also not provoked by reflection. Her discussion of why our consciousness is ambiguous and how the embracement of this ambiguity empowers a person to disclose his being clarifies what Camus identifies as the brief, fleeting fear of the feeling of anxiety that can provoke a person to commit suicide. When a person does not realize the ambiguity of his consciousness, he is trapped by the frustration that ambiguity begets. Furthermore, when a person is trapped in this state, he is blocked from becoming conscious of that feeling of anxiety and ambiguity, and thus can be incited to commit suicide in a moment when he believes that his life is worthless. This belief is marked by a person's diminished self-awareness. The belief that life is meaningless, or absurd, however, is not marked by diminished self-awareness, but

rather by profound conscious awareness. Nishtiani's discussion of how self-questioning inaugurates the religious quest in *Religion and Nothingness* supports this argument, and therefore also serves as support for Camus' argument that conscious awareness of the absurd is triggered by a person questioning his mechanical life, which is an act of self-questioning. Nishtiani's writings also support Camus' argument that this realization of the absurd incites a person to revolt against it and that this revolt enables a person to disclose his being by empowering him to constantly choose not only his actions, but also his own self in spite of the absurdity of his life. In this way, it is evident from the writings of Beauvoir, Nishitani, and Camus that not only is passionate revolt against the absurd the only response to the realization of it, but also that suicide can never be a response to this realization.

References

[1] Simone de Beauvoir, *The Ethics of Ambiguity*, "Ambiguity and Freedom", (NY: The Philosophical Library, 1948)

[2] Albert Camus, *The Myth of Sisyphus*, "An Absurd Reasoning" (NY: Vintage Books, 1955)

[3] Keiji Nishitani, *Religion and Nothingness*, "What is Religion" (CA: University of California Press, Inc.)

BURGE VS. THE LAWYERS

Anonymous

Introduction

Tyler Burge's claim that cogito-like judgments, or second-order judgments, are self-verifying and infallible has generated a great deal of controversy.[55] A second-order judgment is a thought about a thought. This can take the form of propositions such as "I am thinking water is wet." Burge avers that, "these sorts of judgments are self-verifying in an obvious way: making these judgments itself makes them true,"[56] and it does so infallibly.[57] There have been a number of counterarguments to this claim and responses to these counterarguments,[58] but a relatively simple argument against Burge's position has been overlooked, which acts as a defeater and can be made clear by a thought experiment. Before putting forth my counterexample, I shall put Burge's claim into context and briefly explicate what is at stake. After this, I shall present my thought experiment and shall raise and respond to a number of potential objections. Following this, I shall also provide Burge with a position that might save his account, but at the expense of many of his other commitments.

The Context

Tyler Burge's claim of self-verifying and infallible second-order judgments is not a general account of self-knowledge and he does

55. Burge, Tyler. 1988. "Individualism and Self-Knowledge," *The Journal of Philosophy*. Vol. 85, No. 11; pp. 649-663.; Burge,Tyler. 1996. "Our Entitlement to Self-Knowledge," *Proceedings of the Aristotelian Society*. Vol. 96; pp. 91–116.
56. Burge, 1988, 649.
57. Ibid., 658.
58. Since Burge's article first made its appearance in 1988, there have been a number of articles discussing the topic, either with specific devotion to it or as part of a larger point. In 2009 alone, *inter alius*, articles have included: Spicer, Finn. 2009. "On Always Being Right (about What One Is Thinking)," *Canadian Journal of Philosophy*. Vol. 39, No. 1; pp. 137-160; Ren, Huiming. 2009. "Entitlement to Self-Knowledge and Brute Error," *International Journal of Philosophical Studies*. Vol 17, No. 4, pp. 543-562, and several in a special issue of *Erkenntnis* devoted to self-knowledge.

not mean it to be so. Even so, if it is correct, then it entails that self-knowledge is not universally opaque and guarantees that, at least in some instances, self-knowledge is transparent.

In all fairness to Burge, he does not argue for self-verifying self-knowledge for its own sake. Burge's reasoning behind arguing for this type of self-knowledge (in 1988) is to show that self-knowledge is compatible with anti-individualism. Burge also notes that this account is lacking by supplementing it in 1996 with other accounts of entitlement to self-knowledge. My contention is that in describing infallible and self-verifying self-knowledge Burge does not succeed in demonstrating the authenticity of a *certain* (that is, self-referential) type of self-knowledge. This is because the claim is that the knowledge is self-verifying and infallible and this thesis appears to be justified on the basis of the structure of the self-knowledge claims. If I can show instances in which the structure obtains but the self-knowledge does not, then the structure is not sufficient justification for self-knowledge.

Despite Burge not attempting to create a general account of self-knowledge from instances of self-verifying self-knowledge, some seem to have attempted to do just that – or at least a more general account. Sven Bernecker, for example, has criticized Burge's thesis for failing when considering diachronic knowledge and he corrects this defect by discussing first-order to second-order entailment rather than correspondence (a correction that still does not avoid the problems that I raise).[59] In doing so, Bernecker is extending the scope of Burge's thesis and the account of self-verifying self-knowledge is growing more general. Even if Burge himself does not mean for his theory to describe a general account

59. Bernecker, Sven. 2009. "Self-Knowledge and the Bounds of Authenticity," *Erkenntnis*, Vol. 71, No. 1; pp. 107-121. As well, Bernecker takes it to be a "widespread view [that] a second-order judgment of the form 'I believe that p' qualifies as self-knowledge only if the embedded content, p, is of the same type as the content" (107).

of self-knowledge, called by some the "inclusion theory of self-knowledge" – if it is fundamentally flawed, then this flaw ought to be noted before more effort is put forth trying to extend it. For the purposes of this paper, I will treat Burge as the target of my criticisms, as he is the originator of this account of self-knowledge for particular cases, though it is not entirely clear that he ought to be the specific target.

A Thought Experiment

Burge develops his discussion of second-order judgments using the propositions having the form "I am thinking water is wet." Burge claims that if one has such a thought, it must be true – one cannot be mistaken about the corresponding first-order thought in such contexts. To be clear in regards to what Burge means, his claim is that second-order thoughts infallibly lock on to the first order-thoughts. This does not guarantee the truth of first-order thoughts – merely that one is genuinely having a second-order thought that corresponds to the first-order thought. For the sake of clarity, I shall be discussing them as "I am thinking I am thinking water is wet" with that representing the second-order judgment and "'I am thinking water is wet" representing the first-order thought.

The thought experiment proceeds as follows: Imagine there is a man named Marty. Marty is incapable of having a first-order thought that is in violation of his logical schema, but he is not instantaneous in his *recognition of his thoughts*. The logical schema under which Marty operates in the following situation is identical to classical logic.

Marty is called to be part of a jury. During the ensuing trial, the very trustworthy attorney X makes a proposition. The proposition, as with all the others in the trial, are in a type of legalese that Marty believes that he understands but does not in fact understand fully. As a result of the proposition, Marty acquires at t_1 the

second-order judgment, "I am thinking I am thinking [proposition 1]" ($J_m(TP_1)$). Simply because Marty is not completely aware of what attorney X said, it does not mean that he could not have the thought, "I agree with whatever it is that attorney X said, regardless of its semantic content, because he is so trustworthy."

A little later at time t_2, attorney Y, who also appears epistemically trustworthy, makes a proposition 2, which is considered a bromide. He goes on to explain that if one thinks proposition 2, then one must not think proposition 3 ('not think' refers to the lack of assent rather than having a negative thought). Marty, trusting him as well, duly believes proposition 2 and acquires the second-order judgment at t_2, "I am thinking I am not thinking [proposition 3]" ($J_m(\neg TP_3)$).

Subsequently, at t_3, Marty holds the conjunctive second-order thought $J_m(\neg TP_3 \& TP_1)$. Whilst Marty is actively holding this conjunctive thought, attorney X responds by decoding the legalese and informing the jury that P_3 is equivalent to P_1. Until this point, Marty was completely unaware of any relation between the content of P_1 and P_3 or even of the content of either. Marty, being a somewhat conscientious epistemic agent, intentionally replaces P_3 with P_1. So at t_4, Marty replaces $J_m(\neg TP_3)$ with $J_m(\neg TP_1)$, and the philosophically aware Marty, by following the epistemic closure principle[60] and maintaining his adherence to classical logic, holds $J_m(\neg TP_1 \& TP_1)$ simultaneously.

The state of affairs onto which this judgment is supposed to lock is clearly absurd; for it is impossible for someone to hold a thought and simultaneously lack that very thought. Indeed, however, Crispin Wright has noted that the closure principle maintains that knowledge can transfer without the cogency of the argument's needing to be preserved – and it does not mean that the

60. The epistemic closure principle, in its basic form, states that "For all S,ϕ,ψ: if S knows ϕ, and that ϕ entails ψ, then S knows ψ."

person has a "*rational* conviction of the truth of the conclusion."[61] At t_5, Marty eliminates one of the second-order judgments, realizing that he cannot simultaneously hold the thoughts onto which the second-order judgments are intended to lock on to, within his logical schema.

Alternatively, the entire argument can be formulated symbolically:

Let a = proposition 1,
b = proposition 3,
m = Marty,
Bxy = x is thinking y,
Jxyz = x is thinking y is thinking z,
Kxyz = x is thinking y is not thinking z, and
Exxyz = x thinks x thinks y entails z.

1. Jmma
2. Kmmb
3. Emmba
4. $\forall x \forall y \forall z(Exxyz \rightarrow ((Kxxy \rightarrow Kxxz) \& (Jxxy \rightarrow Jxxz)))$[62]
5. $\forall x \forall y(Jxxy \rightarrow Bxy) \& \forall x \forall y(Kxxy \rightarrow \neg Bxy)$[63]
    6. $Jmma \rightarrow Bma$ (From 5)
    7. $Kmma \rightarrow \neg Bma$ (From 5)
    8. Kmma (From 4, 3, 2)
    9. Bma (From 1, 6)
    10. $\neg Bma$ (From 7, 8)
    11. Contradiction (From 9, 10)
12. $\neg(\forall x \forall y(Jxxy \rightarrow Bxy) \& \forall x \forall y(Kxxy \rightarrow \neg Bxy))$[64]

---

61. Wright, Crispin. 2000. "Cogency and Question-Begging: Some Reflections on McKinsey's Paradox and Putnam's Proof," Philosophical Issues, Vol. 10; pp. 140-163. 140. Emphasis added.
62. (Closure) If one wishes to deny closure as a rule but rather accept it as an assumption, then one may change the proof to "If closure, then not Burge's claim."
63. Assumption: Burge's Claim
64. (From 5,11)

(3) and (4) are not required for an anti-Burgean proof to work. (2) can begin as (8) without substitution being required because the contradiction does not arise until Burge's assumption comes into play. The inclusion of (3) and this form of (2) make the situation's occurrence more psychologically acceptable. Without such premises, the example would be open to the objection that no actual person would arbitrarily choose to maintain second-order judgments about contradictory thoughts. This appears to be more problematic than accepting that single-premise closure obtains. Some, however, might deny single-premise closure but think that the psychological rarity of an example in which closure is not imposed is philosophically irrelevant, supporting my conclusion in that manner. Before providing detailed analysis of the situation that has occurred, let us see another example in which (3) and (4) are omitted.

Another Thought Experiment

Imagine that this situation precedes the previous example, and now the logical schema under which Marty operates is that of dialethic paraconsistent logic, and as such, he does not have the principle of non-contradiction as part of his schema.[65] In this example, Marty sees a ball and observes it as being red. At t_1, Marty acquires the second-order judgment, "I am thinking I am thinking the ball is red" (J_m(TBR)). Marty's trusted friend also sees the ball and informs Marty that he observes the ball as a color other than red. At t_2, Marty, trusting his friend's judgment and not believing that contradiction is illogical, decides he ought not to be thinking the ball is red and incorporates this into his beliefs. And so he adds the second-order judgment, "I am thinking I am not

65. One might contend that dialethic paraconsistent logic is not a viable system. While such a system may not be the most appealing or useful one available, with the extensive written work on dialethic paraconsistent logic – particularly by Graham Priest – such a system should no longer thought of as an irrational absurdity.

thinking the ball is red" ($J_m(\neg TBR)$). Thus he has the second-order conjunctive judgment ($J_m(\neg TBR)$ & ($J_m(TBR)$)). At this point, it is *prima facie a reductio* for Marty to hold a conjunctive second-order judgment in which he thinks that he holds a first-order thought-set that is *actually* impossible to hold.

Some, however, may attempt to lodge an objection that due to Marty's holding a system of dialethic paraconsistent logic, his second-order judgments do actually correspond to his first-order thoughts. If this objection does have merit, then the example may be solidified in a preventative way by continuing it in the following manner. Marty then attends a logic class and is informed of classical logic, which he then decides to adopt as his logical schema – replacing paraconsistent logic. So at t_3, his logical schema is identical to classical logic. At t_4, Marty entertains the previously mentioned second-order conjunctive judgment, as the addition of a new schema that supervenes over his old thoughts does not entail an *instantaneous* revision of all his previous thoughts. It seems that he would first genuinely and actively have to hold his previous thoughts before deciding whether to abandon them.

One might object that he would not be *genuinely* actively holding his previous thought on the matter. This, however, is not a satisfactory objection as until this point Marty has no more reason to be suspicious of this particular conjunctive thought than any other. If one argues that Marty would hold these thoughts under duress at the addition of a new schema, then he would have to hold all other thoughts that he previously took for granted under suspicion until he has time to examine them. For that matter, any time that a new thought is added that might have an effect on multiple thoughts, he would have to go through this process, and it would be difficult for Marty, or anyone, to think anything genuinely. At t_5, having realized there is an inconsistency in his thought-set, he eliminates one of the judgments.

Diagnosis and Objections

There are several ways to interpret the data of the scenarios. The first – my claim – can be stated as follows: the problem for Burge occurs at t_4. At t_4, Marty holds two (supposedly infallible) second-order judgments simultaneously, which refer to two first-order thoughts that cannot exist together in a logically consistent schema that includes the principle of non-contradiction. For one cannot both hold a thought and not hold a thought – though one can hold a thought that P and a thought that not-P as the holding of contradictory thought is not an extensional contradiction but only intensional (allowing for the plausible presence of the second-order thought about contradictory first-order thoughts in Marty's mind). So it must be the case that he does not actually think one of the first-order propositions, or perhaps either of them, and is mistaken in his judgment. Therefore it is the case that the second-order judgment is not infallible.

A different contention is that Burge's claim does not entail that the second-order thought about the lack of thought entails the first-order lack of thought. Burge's claim only refers to positive thoughts claims, and this is a negative belief. By a positive thought, we mean a thought that need not refer to a negation of thought P. By a negative thought, we mean a thought that invokes a negation of a thought P. This objection does not appear to be plausible. Not only does Burge not specify this, but the distinction between "positive" and "negative" does not seem to be relevant to his thesis. He never discusses the point and it would be *ad hoc* for Burge to claim that of all possible introspective thought, the only ones that are opaque are thoughts about suspended thoughts – a specific type of thought-lacking. I object that it would be *ad hoc* because there is no principled reason to deny the entailment, independent of my argument requiring a response. A distinction between positive and negative thoughts is irrelevant because the lack of thought can be reformed without losing any meaning as an

affirmation of the suspension of thought. Imagine a person thinks "I am thinking water is wet" and "I am suspending judgment about water's being wet". According to a straightforward reading of Burge, both second-order thoughts entail the corresponding first-order thoughts.

One might also object that it is actually impossible for someone knowingly to hold contradictory thoughts simultaneously in one's mind, and so the Marty case, while a logically possible situation, is not a metaphysically possible one. The instant that one is made aware of a contradiction, at least one of the thoughts is eliminated or suspended. The phenomenon that appears with Marty, however, seems *prima facie* to reflect commonplace events that most people would accede to having experienced. That is, one has two thoughts and one is informed of a contradiction, and one takes time to ponder over what one thinks. Despite knowing that one cannot think both thoughts at the same time and be logically consistent, one needs to process the information. It would not be held as unusual to hear someone say upon discovery of such a conflict, "I think x, but I also don't think y [which is equivalent to x, or entails x]; how do I reconcile this?" This would, in fact, seem more likely than one's instantly saying that one must immediately correct one's thoughts. Everyday examples of this might include political positions. One might suppose that a person thinks that they hold the party line in all positions (and is aware of these positions) but she also has an aberrant thought about a certain political issue. In this, her thought about issue X is that she must have more time to cogitate about it and in the meantime must suspend judgment. This suspension of judgment is *contra* party line and she acknowledges this fact but needs time to resolve the conflict. In the mean time, she holds that she both maintains the party line and thinks a position that entails a deviation from the party line.

There is one more point that must be addressed before I may

continue. Oftentimes, philosophers appeal to varying degrees of confidence in knowledge to account for certain problems in epistemology. I have attempted to phrase the previous examples in such a way that an appeal of this sort as a mode of objection against the plausibility of my examples cannot be made. If, however, I have overlooked something, I should note that the means of acquisition of such second-order judgments is irrelevant. What is important is that the state of affairs that arises at t_4 is logically possible. It could have very well arisen due to a mad scientist implanting such judgments in Marty. The scientist could, indeed, have implanted the second-order judgment into Marty, "I am thinking I am not thinking water is wet" while still leaving all of Marty's previous experiences with water intact (in which Marty has experienced that water is in fact wet). If this is the case, the scientist would have implanted the second-order judgments in Marty with full confidence, and t_4 would still come to be.

There is another way to interpret the data from which the title of the paper is derived. It could be that Marty, in ever having these second-order thoughts about contradictory first-order thoughts, is committing an impossible act if Burge is correct (and yet Marty is doing so), regardless of whether they accurately lock on to the first-order thoughts. The reason for this is that if they are both infallible in their correspondence to the first-order thoughts – and one has direct access to these thoughts – then any logical person operating within the schema that maintains the principle of non-contradiction immediately would have to reject these self-verifying thoughts due to their irreconcilability. Marty cannot have contradictory self-verifying thoughts if the principle of non-contradiction holds – but he does. Burge may need to claim that the principle of non-contradiction is false and that some version of dialetheism is correct. If this is the case, then he may simply claim that holding $J_m(\neg TBR)$ is compatible with holding $J_m(TBR)$ and Marty would maintain the contradictory and corresponding

first-order thoughts that are entailed under his view. Dialetheism, however, does not appear to be consistent with Burge's writings.

Concluding Remarks

There are many other conceivable scenarios in which this problem might arise. As a result of this, it has been demonstrated that second-order judgments are not infallible or that Burge is a dialetheist. Perhaps Burge's thesis of infallible self-verifying self-knowledge is salvageable, but if so then a novel approach is needed to succeed.