



UNIVERSITY OF EDINBURGH PHILOSOPHY SOCIETY JOURNAL

Undergraduate Philosophical Papers published
for a general academic audience.

ISSUE 1: WINTER 2012
LIST OF CONTRIBUTORS:

1. CHARLOTTE COURSIER
'MODAL THEORIES OF KNOWLEDGE'
2. WILLIAM HOGARTH
'RATIONALITY OR RICHES?'
3. JONATHAN PUGH
'DOES CRAIG'S KALAM ARGUMENT PROVIDE A CONVINCING
ARGUMENT FOR THE EXISTENCE OF GOD?'

EDITED BY: CHRISTOPHER RANALLI AND RICHARD S. GOW

TABLE OF CONTENTS

1. EDITORIAL INTRODUCTION BY CHRIS RANALLI	PG. 3-4
2. ‘MODAL THEORIES OF KNOWLEDGE’ BY CHARLOTTLE COURSIER	PG. 5-17
3. ‘RATIONALITY OR RICHES?’ BY WILLIAM HOGARTH	PG. 18-24
4. ‘DOES CRAIG’S KALAM ARGUMENT PROVIDE A CONVINCING ARGUMENT FOR THE EXISTENCE OF GOD?’ BY JONATHAN PUGH	PG. 25-39
5. EDITORIAL AFTERWORD BY RICHARD S. GOW	PG. 40-41

EDITORIAL INTRODUCTION

The *University of Edinburgh Philosophy Society* dates back to 1871. Since then it has provided a venue for excellent philosophical discussion and debate. But what it has not offered—at least, not until now—is a venue where students and members of the *Society* could exchange philosophical ideas through the medium of written word.

The aim of the Journal is simple and the scope of the Journal is broad. The aim is simple because the Journal is intended to be the written home of the members of the *University of Edinburgh Philosophy Society* and a place where student discussion and debate can both proceed outside the usual venues—e.g., the lecture theatre, classroom, and pub—and into print. But it is also a place for our members to engage with philosophers and students from outside of the *University of Edinburgh Philosophy Society*. Likewise, the scope of the Journal is broad. We welcome—indeed, *encourage*—papers from among the range of philosophical logic, normative ethics, metaphilosophy, aesthetics, philosophy of mind, metaphysics, the relation between biology and philosophy, post-modernism, epistemology, pop-culture and philosophy, and more.

Let me pause to explain the format of each issue. Each issue contains a short *Introduction* to the contents of the current issue and a short *Afterword* which seeks to introduce topics for the next issue. Our first issue features articles from two outstanding graduates of the University of Edinburgh. The first paper, by Charlotte Coursier, titled ‘Modal Theories of Knowledge’, argues that certain modal conditions on propositional knowledge—conditions such as such as the *Safety Condition*, championed by Ernest Sosa and our own Duncan Pritchard, and the *Sensitivity Condition*, which famously figures in the work of Robert Nozick’s (1981) *Philosophical Explanations*—do not provide an adequate explanation of what knowledge is. The second paper, by William Hogarth, focuses on Newcomb’s paradox, and argues that, contra Eells’ (1982), that Evidential Decision Theory does *not* recommend a two-box solution. The third paper, by Jonathan Pugh, titled ‘Does Craig’s *Kalam* Argument Provide a Convincing Argument for the Existence of God?’ argues that William Craig’s *Kalam* argument—a type of *cosmological argument*—fails to provide a convincing argument for the existence of God.

Finally, we encourage potential authors to submit short responses to the papers featured in our previous issues. What we want is for students from within the University of Edinburgh, members of the *University of Edinburgh Philosophy Society*, and students from outside the University of Edinburgh, to engage in lively written philosophical debate and discussion. Indeed, we encourage co-authored papers arguing for a particular philosophical thesis, or co-authored papers responding to previously published articles in our journal. Likewise, attendees of our *Lecture Series* are welcome to pursue further debate in our journal. Perhaps disputes have arisen as a result of a guest lecturer's topic and perhaps you feel that you want to pursue your concerns further. Again, if this is the case—and we hope that it is—then feel free to submit a short piece to our journal for review.

Enjoy the *University of Edinburgh Philosophy Society's* first Journal issue! We look forward to your submissions and comments.

Sincerely,

Chris Ranalli
Richard S. Gow
Editors

MODAL THEORIES OF KNOWLEDGE

CHARLOTTE COURSIER

1. Introduction: The problem of analysing knowledge

Prior to Gettier's paper 'Is Justified True Belief Knowledge?' (1963), the tripartite definition was widely held to be an adequate account of knowledge.¹ Gettier presented his argument in the form of counterexamples where an agent has justified true belief but does not have knowledge. The key to Gettier's, and other similar 'Gettier', counterexamples is that the belief in question is obtained or made true by some sort of 'epistemic luck'. Specifically, the difficulty is that possessing justification for true belief doesn't eliminate that belief's being lucky. However, only some forms of luck are epistemically relevant. To restrict the relevant kinds of luck there are various proposed conditions: for example, to require a causal connection (Goldman 1967), or that the belief be produced by a reliable method or process (Goldman 1979), or by virtuous cognition (Sosa 2003, Zagzebski 2001). Here I will concentrate on explaining luck in terms of modality. The idea is to consider how the facts concerning the agent's belief vary with the truth value of the proposition across possible worlds. There are three conditions: sensitivity (Dretzke 1970, Nozick 1981), adherence (Nozick 1981) and safety (Sosa 1999, Williamson 2000, Pritchard 2005).² I shall elaborate on epistemic luck, the modal conditions and their motivations. I shall then compare and evaluate the conditions in light of how they perform with respect to Gettier-style counterexamples; concluding that, whilst safety seems to be the strongest of the three, the modal theories are not adequate theories of knowledge.

2. Modal Conditions on Knowledge

¹ Plato, *Theaetetus*, 201d: 'knowledge is true judgement with an account'.

² Dretzke (1970) alludes towards a condition like sensitivity, but it is first overtly described by Nozick (1981). Although Nozick originally calls the condition 'variation' I shall refer to it as 'sensitivity' to be consistent with the literature (Nozick 1981: 172-196). Similarly, 'adherence' is not Nozick's own terminology (Lacey 2001: 103).

2.1 *The conditions*: the motivating idea behind modal conditions is that ‘epistemic luck’ can be explained as a coincidence between the belief and the proposition’s being true; had things been slightly different, the coincidence wouldn’t have obtained. Thus, they involve a subjunctive conditional which restricts knowledge with respect to what could have been the case in nearby worlds. That is, they try to ensure that the belief in a certain proposition ‘tracks the truth’ of that proposition across possible worlds.³ These conditions resolve many of the previous Gettier-style counterexamples.⁴ Three such modal conditions have been proposed:

Sensitivity:

S’s belief that p is *sensitive* iff the nearest $\neg p$ -worlds are $\neg Bp$ -worlds ($\neg p \square \rightarrow \neg Bp$).

Adherence:

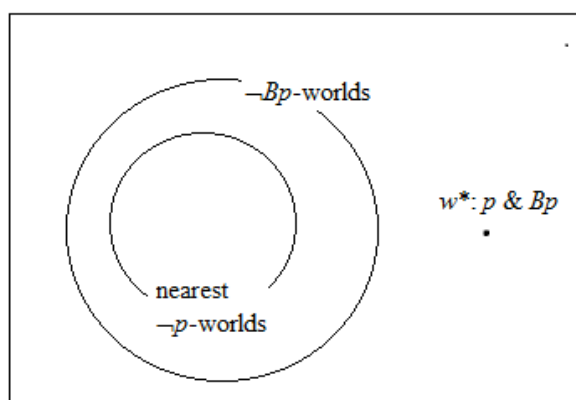
S’s belief that p is *adherent* iff the nearest p -worlds are Bp -worlds ($p \square \rightarrow Bp$).

Safety:

S’s belief that p is *safe* iff the nearest Bp -worlds are p -worlds ($Bp \square \rightarrow p$).⁵

One could represent these three modal conditions diagrammatically as follows (w^* is the actual world, in which p & Bp are true):

Figure 1: **Sensitivity**



³ Modal theories are externalist positions because they deny that if an agent knows p then they know that they know p (the ‘KK principle’).

⁴ Such as Gettier’s original examples, the stopped clock, fake barns and fake sheep.

⁵ I am using Holton’s explanation of the truth conditions for subjunctive conditionals (see Holton 2004). These truth conditions come from Lewis’s *Counterfactuals* 1973. The original theory of possible worlds comes from Leibniz’s theory of possible worlds in *Theodicy* 1709.

Figure 2: Adherence

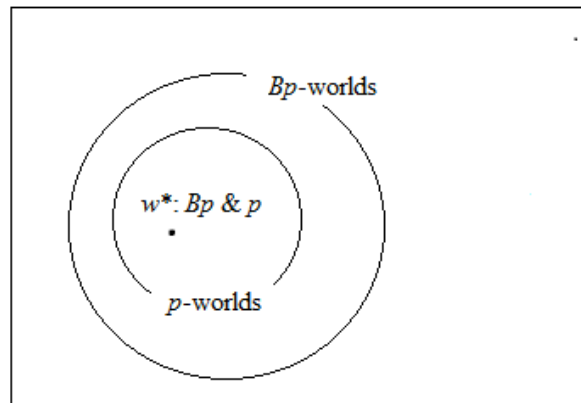
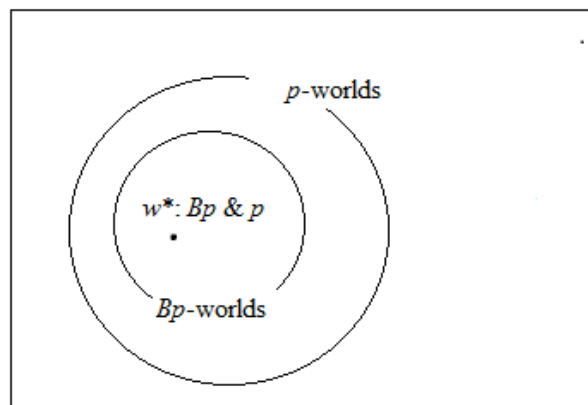


Figure 3: Safety



Subjunctive (or modal) conditionals, unlike material conditionals, don't validly contrapose. That is, a material conditional of the form $p \rightarrow q$ is equivalent to $\neg q \rightarrow \neg p$. However, for subjunctive conditionals this does not hold. In particular, $Bp \Box \rightarrow p$ (i.e., safety) is not equivalent to $\neg p \Box \rightarrow \neg Bp$ (sensitivity). This is because even if the nearest Bp -world is a p -world, the nearest $\neg p$ -world might also be a Bp -world. For example, an agent truly believes 'I am not a brain in a vat'. In nearby worlds where the agent also believes this, it is still true; so this true belief is safe. But in the nearest world where the agent *is* a brain in a vat, the agent still believes he isn't. So, the true belief is not sensitive. This is because worlds where

a sceptical hypothesis is true are far off. However, sensitivity, in this case, looks at the nearest worlds where the sceptical hypothesis is true.

2.2 *Motivations*: the modal conditions resolve many of the initial Gettier-style cases. For example, an agent looks at a stopped clock which reads the time to be 8 o'clock. It just so happens to be 8 o'clock. So, the agent has a true (because it is 8 o'clock) justified (because usually using a clock is an adequate justification for telling the time) belief. However, it is lucky that the agent's justified belief is true. If the agent had looked at the clock a minute earlier or later, or if the clock had stopped at a slightly different time, then they would have formed a false belief. The modal conditions restrict this form of luck. If it wasn't the case that it was 8 o'clock the agent would still believe that it was 8 o'clock (i.e., $\neg(\neg p \Box \rightarrow \neg Bp)$). Therefore, the belief is not sensitive and thus not knowledge according to sensitivity. If the clock had stopped at a slightly different time then the agent would no longer believe it to be 8 o'clock (i.e., $\neg(p \Box \rightarrow Bp)$). Therefore, the belief is not adherent and thus not knowledge according to adherence. Finally, if the agent had looked at the clock at a slightly different time they would still believe it to be 8 o'clock but this belief would be false (i.e., $\neg(Bp \Box \rightarrow p)$). Therefore, the belief is not safe and thus not knowledge according to safety. From this example it is clear to see how the three modal conditions, in slightly different ways, restrict the luck that prevents the agent having knowledge, despite their possessing a justified true belief.

One particular motivation for the sensitivity condition is that it preserves common sense knowledge, while accepting that we cannot know the denials of sceptical hypotheses. For example, if an agent believes that they are not a brain in a vat ($\neg BIV$), were this belief false, they would continue to believe it (i.e., $\neg(p \Box \rightarrow Bp)$). However, typical common sense knowledge, such as, 'I have hands', is sensitive. This is because in the *nearest* worlds where \neg 'I have hands', I no longer believe that 'I have hands'. Despite, BIV implying that \neg 'I have hands'. Thus, it denies the closure principle (i.e., $[Kp \ \& \ K(p \rightarrow q)] \rightarrow Kq$). The safety condition wants to preserve the closure principle. Furthermore, safety seems to effectively deal with counterexamples to sensitivity's necessity. Finally, the adherence condition is motivated by its ability to deal with knowledge of necessary truths, which shall be discussed below.

2.3 *Epistemic luck*: a lucky event is often characterized by the low probability of its occurrence. However, while this is a necessary condition two further

components need to be outlined, in contrast to concepts such as accidental or chance events, a lucky event: (1) follows from the initial conditions in the actual world; and (2) is significant to an agent *S*. Pritchard (2005) offers two conditions of luck with this in mind. Firstly, in terms of counterfactuals:

(*L*₁) If an event is lucky, then it is an event that occurs in the actual world but which does not occur in a wide class of the nearest possible worlds where the relevant initial conditions for that event are the same as in the actual world (Pritchard 2005, 128).

Secondly, in terms of the agent:

(*L*₂) If an event is lucky, then it is an event that is significant to the agent concerned (or would be significant, were the agent to be availed of relevant facts) (Pritchard 2005, 132).

Furthermore, there are various types of luck that could be involved in the formation of a belief. Not all of these types of luck hinder knowledge. The modal conditions must be set up so they ensure that they *only* eliminate 'epistemic luck'. Unger (1968, 159) outlines two types of luck that are unproblematic: (1) luck that *p* is true; and (2) luck that *S* has the capacity for knowledge (e.g., has the cognitive ability or exists at all). Pritchard (2005, 136-8) offers two further types: (3) luck that *S* acquires their justification to *Bp*; and (4) luck that *S Bp*. The type of epistemic luck relevant to Gettier cases is in terms of how the agent *in fact* formed their belief:

(*EL*₁) luck that, given the facts external to the agent in the actual world, the agent's belief is true.

A second type of epistemic luck, as outlined by Pritchard (2005, 175), concerns what the agent *knows*, or is able to *know*, about how they formed their belief:

(*EL*₂) luck that, given what the agent is able to know by reflection alone in the actual world, the agent's belief is true.

Both of these forms of epistemic luck concern the relationship between the agent's justification for *Bp* and the fact that *p* is true. However, *EL*₁ outlines it in terms of the external facts (such as, what time the clock stops); while *EL*₂, outlines

it in terms of the facts known (or knowable) by the agent about their justification to believe p (such as, their beliefs about how reliable the clock is).⁶

3. Adequacy Problems

3.1 *Sensitivity, adherence and safety are not jointly sufficient for knowledge*: Bill contracts an illness that renders him red-green colour blind. Because of this he stays in bed while he is sick. Nothing in his room is red or green. So, he doesn't know that he is temporarily red-green colour blind. By coincidence, the glass in his window reverses the effects of red-green colour-blindness, but has no effect on normal perceivers. Therefore, Bill doesn't know the glass has this effect. So, when Bill looks out the window and sees his nephew wearing a green t-shirt, he believes, truly, that his nephew's t-shirt is green. However, intuitively this is not a case of knowledge. It's lucky that the glass happens to cancel colour-blindness. However, his belief is safe because in nearby possible worlds where he believes his nephew's t-shirt is green it actually is green. Bill's belief is also sensitive because in the nearest world where his nephew's t-shirt isn't green, Bill doesn't believe that it is green. Finally, Bill's belief is adherent, because in nearby worlds where his nephew's t-shirt is green, Bill still believes that it is green. What makes this case distinctive, and others like it, is that the lucky component that is preventing the agent from acquiring knowledge, rather than simply a justified true belief, is enabling the agent to some degree reliably track the truth of the proposition. However, we are disinclined to ascribe knowledge to the agent because they are failing to exercise a cognitive ability (i.e., they lack reliable colour perception).⁷

3.2 *Sensitivity is not necessary for knowledge*: I believe, truly, that there is coca cola in my coke can, based on the appearance of the can. Intuitively, I know that there is coca cola in my coke can. However, in some nearby world there was a mistake at the factory, resulting in there being sprite in my coke can. But in this world I still believe that there is coke in it. Therefore, it is not the case that $\neg p \square \rightarrow \neg Bp$. That is, my true belief is not sensitive even though it is a case of

⁶ This particular type of epistemic luck is relevant to why it seems to be necessary to include an internalist account of methods in a modal condition on knowledge (a modification of the principles I discuss below).

⁷ Furthermore, if we want to incorporate an internalist account of methods into the conditions (which will be outlined below) then the agent lacks knowledge of why their method is justified (i.e., they don't know that (1) they are temporarily colour blind; and (2) they don't know that the glass is cancelling this to make their beliefs 'track the truth' of the proposition).

knowledge. This case shows that the sensitivity condition is too strong: there are many similar cases of common sense that we want to preserve in our account of knowledge.⁸ A defender of the sensitivity condition could argue that this just simply isn't a case of knowledge. However, this seems contrary to one of their own motivations: to preserve cases of common sense knowledge like this. Moreover, the sensitivity condition analyses the closest worlds where p is false – no matter how far off from the actual world they are. However, the *reason* that the appearance of the coke can is an adequate justification for knowledge that 'there is coke in it' is precisely *because* $\neg p$ worlds are far off (i.e., it is highly unlikely that a mistake in the factory would occur, due to all the rigorous safety protocols ...etc). Sensitivity fails to capture these kinds of justifications. It is clear from cases like this how safety and sensitivity are not equivalent (because they don't validly contrapose): all the nearest worlds which are Bp -worlds, are also p -worlds, so the belief is safe.

3.3 *Adherence is not necessary for knowledge*: I believe, truly, that my friend is walking down Princes St because I just saw her walk past the coffee shop window. Intuitively this is a case of knowledge. However, I was reading the paper and just happened to look up by chance at the moment she walked past. This true belief is not adherent because in a nearby world where I had not looked up from my paper at that moment I would not know that my friend was walking down Princes St. That is, some nearby p -world is not a Bp -world. Therefore, one can know a proposition, though the true belief is not adherent; so, adherence is not necessary for knowledge. This case shows that the adherence condition is too broad: it attempts to eliminate epistemic luck at the cost of eliminating justification obtained by chance.

3.4 *Safety is not necessary for knowledge*: the police are after Morrissey because they believe he is planning to export border restricted anti-depressants to New York from the United Kingdom. They know that he plans to catch the 13.30 flight from Edinburgh Airport. They tell the airport staff that if Morrissey arrives they should tell him that the flight is leaving from Gate G and then call the pilot to move the aeroplane to Gate A. Mark, who likes to dress like Morrissey because he is a big fan, is also planning to catch the same flight. However, at the last minute he decides to dress down for comfort. So, when he

⁸ Sosa's famous rubbish chute example makes the same point: the sensitivity condition expects too much (see Sosa 2000, 13-14).

asks the airport staff which Gate the 13.30 New York flight is at they tell him Gate G. He then believes, truly because Morrissey fails to arrive, the flight is at Gate G. Intuitively, it seems that Mark knows that the flight will be at Gate G, as he has sufficient justification from the airport staff. However, his belief could easily have been false. In a nearby world, he decides to dress like Morrissey but he still believes that the flight leaves from Gate G. But then, this belief is false. That is, he has knowledge but it is not safe.⁹ Note that the luck involved in this case does not affect the cognitive ability of Mark. Thus, the luck is not epistemically relevant and therefore, despite Mark's belief not being safe it is still an instance of knowledge (and therefore a counterexample to safety's necessity).

3.5 *Sensitivity and safety are not jointly sufficient for knowledge of necessary truths:* the existence of God is either necessarily true or necessarily false. So, if God exists but Oscar is a theist for bad reasons (e.g., he mistakenly believes he's seen a miracle – he believes a statue of the Virgin Mary was crying when there was just a leak in the roof), his belief is still, safe and sensitive since there are no possible worlds where God does not exist. And, similarly, if God doesn't exist but Hannah is an atheist for bad reasons (e.g., she likes Christopher Hitchens), her belief is still safe and sensitive. This is because necessary truths will remain the same in all possible worlds. However, intuitively neither Oscar nor Hannah have knowledge. Analogously, this also holds for logical, mathematical and scientific necessary truths. Neither safety nor sensitivity are sufficient for knowledge of necessary truths. However, adherence is sufficient for knowledge of necessary truths: since both Oscar and Hannah could easily not have formed their beliefs, while p 's truth value would remain constant.

3.6 *Closure:* Nick knows that he has hands; he also knows that if he has hands then he is not a brain in a vat (\neg BIV). He also believes \neg BIV. According to closure (i.e., $[Kp \ \& \ K(p \rightarrow q)] \rightarrow Kq$), he must also know \neg BIV. However, Nozick insists that he doesn't know \neg BIV, because his belief that \neg BIV is not sensitive. Nozick claims that this insensitivity explains why he doesn't know \neg BIV. So, on Nozick's modal account, the closure principle is violated. Although Nick's belief that \neg BIV is not sensitive, it is safe (in close possible worlds where he believes \neg BIV, his belief is true). However, the consequences of safe true beliefs are also sometimes not safe. Murphy (2005) describes a case adapted from Kripke's

⁹ This is a modification of the 'Halloween Party' case presented by Comesaña (2005, 397).

original 'red barn' example, where Harriet sees a red barn in a field.¹⁰ There are also many other fake non-red barns in the field. Harriet believes, truly, that there is a red barn in the field. This belief is safe because there are no fake red barns in the field. Her belief could not easily have been false. From her belief she infers that there is a barn in the field. So, she believes, truly, that there is a barn in the field. However, this is not a safe belief because in a nearby possible world she could have looked at one of the fake barns instead.

4. Modifications

Many defenders of the various modal conditions acknowledge that they are not sufficient for knowledge and so add further conditions.¹¹ However, as they have been shown to be unnecessary for knowledge then ultimately this won't help. Defenders have proposed augmenting the conditions with respect to the method that the belief was originally obtained. Firstly, with adherence:

Necessarily, if *S* knows that *p* via method *M* then, if *p* is true, *S* continues to believe *p* via *M*.

Nozick outlines an internalist conception of methods: *S* knows why their method for believing *p* is justified. However, this modification does rule out the above counterexample (the Princes St case). This is because, in the close possible worlds where the agent does not look up and see their friend, they are no longer using the same method as in the actual world (i.e., the use of the agent's perceptual sight). However, there are further difficulties for the adherence condition. It implies that *S* needs to always *Bp* to be the case via the same method in all close worlds where *p* continues to be true. However, in some close possible worlds *S* could be presented with further evidence that makes *S* $\neg Bp$ despite it not affecting the validity of their justification. For example, *S* believes that Tesco is open, based on the justification that they regularly attend the shop and knows their opening hours well and has phoned up to double check. And it's true, Tesco is open. However, in a nearby possible world *S* happens to overhear someone saying that Tesco is closed; *S* now $\neg Bp$, despite *p* still being the case. They simply with-hold judgement, and decide to go to the shop in the other direction to save the possible extra walk. It may be the case that *S* is over-

¹⁰ Although, Kripke gave the example in a conference in the 1980s, it was never published.

¹¹ For example, both Pritchard and Sosa also endorse virtue epistemological conditions (see Pritchard 2009 and Sosa 2003).

cautious in their belief formation, but it doesn't seem clear that this implies they lack knowledge in the actual world. Someone with lower standards for their belief formation, according to adherence, would have knowledge in this case, despite *S* failing to (Bird 2003, 9-10).

Secondly, with sensitivity:

Necessarily, if *S* knows that *p* via method *M* then if *p* were false, *S* would not believe *p* via *M*.

That is, if the agent *were* to use the same method in the counterfactual scenario, where the proposition *p* is false, they would not believe *p*.¹² For example, my *method* (or justification) for believing that there is coca cola in my coke can is that the can is the familiar coke can that has always contained coca cola.¹³ This modification of the principle works for some cases, such as the grandmother case that Nozick has in mind: originally the grandmother believes that her grandson is well based on seeing that he looks well and it's true, he is well (Nozick 1981, 179). However, in a possible world where her grandson is ill, the family would tell her that he's well to save her worrying. So, she would still believe him to be well even though the proposition is now false. This is a case of knowledge that is not sensitive. However, in the counterfactual scenario the grandmother's method is different: originally her method is seeing that her grandson is well, in the counterfactual scenario it is the testimony of her family. So, according to the new principle the belief is sensitive because if the grandmother *were* to use the same method in this possible world she would no longer believe her grandson to be well.

Finally, with safety:

S's belief is safe if in most near-by possible worlds in which *S* continues to form her belief about the target proposition *in the same way* as in the actual world, and in all very close near-by possible worlds in which *S* continues to

¹² This is a modification of Nozick's original formulation that the agent must use their *actual* method in the possible world scenario (see Nozick 1981, 184-5). Becker suggests that using the same method in a counterfactual scenario entails there being the same evidence (from a criticism raised by Luper-Foy), so it must be that if the agent *were* to use the same method, and the proposition is false, they would not believe *p* (see Luper-Foy 1984, 28-9 and Becker 2009, 20)

¹³ Note that the method is not a proposition, so it being internal does not conflict with Nozick's content externalism (i.e. an agent may not know that they are using a certain method) (see Becker 2009, 21-26).

form her belief about the target proposition *in the same way* as the actual world, her belief continues to be true.¹⁴

That is, if the agent were to form her belief of proposition p via the same method in most counterfactual scenarios then p would continue to be true. In a similar respect the modified safety condition deals with a range of counterexamples to its necessity where the same method is not being deployed in the counterfactual scenario. However, in both of the counterexamples presented above (i.e. the coke can and Morrissey examples) the methods in the counterfactual scenarios are the same (i.e. the appearance of the coke can and the testimony of the airport staff respectively). This example demonstrates that just because a method may be unreliable in a nearby world it does not necessarily imply that it is unreliable in the actual world.

5. Conclusion

None of the modal conditions seem to give a satisfactory account of knowledge. There are counterexamples which show all three conditions to be jointly insufficient. Moreover, by their very nature modal conditions cannot account for knowledge of necessary truths as their truth values remain the same in all possible worlds. Adherence is proposed by Nozick to deal with this difficulty. However, it seems adherence is an unnecessary condition because it implies two agents can differ as to whether they have knowledge in the same circumstances. Furthermore, both safety and sensitivity conflict with the plausible closure principle. Adding further conditions would resolve problems of sufficiency. Most problematic however, is that even once the conditions have been augmented with an inclusion of the agent's method, neither safety nor sensitivity seem to be necessary for knowledge. Sensitivity seems to be too strong: it rules out cases of common sense knowledge based on inductive inference. Safety does seem to be the most plausible of the modal conditions, but it still requires an adequate explanation of how it seems to fail to be necessary in the case presented above. Adding conditions won't help. A further modification of the principle, over and above an inclusion of the method used by the agent, seems to be the most promising approach.

¹⁴ This is a modification of a principle originally defended in Pritchard (2007, 282) and then presented *with* modifications in Pritchard (2009, 34). Emphasis added.

References

Becker, K. 2009: 'Margins for Error and Sensitivity: What Nozick Might Have Said' *Acta Analytica* **24**: 17-31.

Bird, A. 2003: 'Nozick's Fourth Condition' *Fact Philosophica* **4**: 141-51.

Burnyeat, M. 1990: *The Theaetetus of Plato*. J. Levett (Trans.) Hackett: Indianapolis.

Comesaña, J. 2005: 'Unsafe Knowledge' *Synthese* **146**: 495-404.

Dretzke, F. 1970: 'Epistemic Operators' *Journal of Philosophy* **67**: 1007-23.

Gettier, E. 1963: 'Is Justified True Belief Knowledge?' *Analysis* **23**: 121-123.

Goldman, A. 1967: 'A Causal Theory of Knowing' *Journal of Philosophy* **64**: 357-372.

Goldman, A. 1979: 'What is Justified Belief' in G. Pappas (ed.) 1979, *Justification and Knowledge*. Dordrecht: Reidel. pp. 1-23.

Holton, R. 2004: 'Counterfactuals' *Lecture notes*. Massachusetts Institute of Technology.

Kallestrup, J. 2010: 'Are We Safe from Sensitivity?' *Manuscript*. University of Edinburgh.

Lacey, A. 2001: *Robert Nozick*. Princeton: Princeton UP.

Leibniz, G. 1709: *Theodicy*. A. Farrer (ed.) and E. Huggard (Trans.) New Haven: Yale UP. (1952).

Lewis, D. 1973: *Counterfactuals*. Oxford: Blackwell.

Luper-Foy, S. 1984: 'The Epistemic Predicament: Knowledge, Nozickian Tracking, and Skepticism' *Australian Journal of Philosophy* **62** (1): 26-49.

Murphy, P. 2005: 'Closure Failures for Safety' *Philosophia* **33**: 331-334.

Nozick, R. 1981: 'Knowledge and Scepticism' in *Philosophical Explanations*. Cambridge, Ma.: Harvard UP.

Nudds, M. 2008: 'Truth-Tracking' *Lecture notes*. University of Edinburgh.

Pritchard, D. 2005: *Epistemic Luck*. Oxford: Oxford UP.

Pritchard, D. 2007: 'Anti-Luck Epistemology' *Synthese* **158** (3): 277-97.

Pritchard, D. 2008: 'Sensitivity, Safety, and Anti-Luck Epistemology' in J. Greco (ed.) *The Oxford Handbook of Skepticism*. Oxford: Oxford UP. pp. 437-55.

Pritchard, D. 2009: 'Safety-Based Epistemology: Whither Now?' *Journal of Philosophical Research* **34**: 33-45.

Pritchard, D. 2010: *What is This Thing Called Knowledge?* 2nd edition. London: Routledge.

Sosa, E. 1999: 'How Must Knowledge Be Modally Related to What Is Known?' *Philosophical Topics* **26**: 373-384.

Sosa, E. 2000: 'Skepticism and Contextualism' *Philosophical Issues* **34**: 1-18.

Sosa, E. 2003: 'Virtue Epistemology' in *Virtue Epistemology: Internalism Versus Externalism*. Oxford: Blackwell.

Unger, P. 1968: 'An Analysis of Factual Knowledge' *Journal of Philosophy* **65**: 157-70.

Williamson, T. 2000: *Knowledge and Its Limits*. New York: Oxford UP.

Zagzebski, L. and Fairweather, A. 2001: *Virtue Epistemology: Essays on Epistemic Virtue and Responsibility*. Oxford: Oxford UP.

CAUSAL DECISION THEORY OR EVIDENTIAL DECISION THEORY: RATIONALITY OR RICHES.

WILLIAM HOGARTH

The respective merits of Causal Decision Theory's (CDT) and Evidential Decision Theory's (EDT) models of rationality can be made clear by a comparison of their solution to Newcomb's problem. First, I will argue that it is rational to take both boxes in Newcomb's problem. Secondly, I shall argue that CDT recommends two-boxing but that EDT does not. Specifically, I shall argue that Eells's claim the EDT recommends two-boxing is false, and that this is because his 'tickle defence' of EDT fails. I will conclude that CDT provides a better model of rationality than EDT because CDT is able to distinguish auspiciousness from efficacy.

Newcomb's problem, also known as the predictor paradox, is a thought experiment designed to test our intuitive conception of rationality. Imagine that there is a being called the Predictor who is able to predict the choices that you make with almost complete accuracy. You must now play a game with the Predictor. The Predictor will present you with two boxes, one opaque and one transparent. You are given a choice; you can either take both boxes, or you can take just the opaque box. Inside the transparent box there is \$1,000. The content of the opaque box is determined by the choice that the Predictor thinks you will make: if he thinks you will take both boxes the opaque box will be empty, if he thinks you will only take the opaque box it will contain \$1,000,000. When you come to make your choice the content of the boxes cannot be changed; the \$1,000,000 is either in the opaque box or it is not and this cannot be changed at this point in the game. Given the high probability (> 0.9) that the Predictor accurately predicts your choice, one-boxing seems to lead to riches; you'll get your \$1,000,000. I will now argue that two-boxing is the rational strategy to employ when facing Newcomb's problem.

When faced with Newcomb's problem two-boxing is rational because it dominates one-boxing. First, assume that the predictor has predicted that some agent *A* will take just the box with the \$1,000,000 in it. Assuming that expected monetary value is equal to utility then *A* will maximise utility by taking both boxes because *A* will get the \$1,000 as well. Secondly, assume that the predictor

has predicted that *A* will take both boxes, and so the \$1,000,000 is not in the second box. Taking both boxes will still maximise utility because *A* will at least get the \$1,000. Therefore, two-boxing dominates one-boxing. The reason that two-boxing dominates one-boxing is that the features related to the utility of possible outcomes are causally independent of the agent's actions (Sobel, 1994, p32). However, these features are evidentially dependent on the agent's actions (Sobel, 1994, p32). This leads some to think that one-boxing is the rational strategy, but this is to fail to distinguish between efficacy and auspiciousness (Weirich, 2008): The decision to take one box is auspicious of there being the \$1,000,000 in the box but has no efficacy in bringing it about. Once the agent is in a position to make their choice, their choice cannot causally affect what is in the two boxes. Therefore, I conclude that it is rational to take two boxes and that a failure to take two boxes results from a failure to distinguish efficacy from auspiciousness. Thus, if it cannot be argued that EDT advocates the two-boxing strategy when faced with Newcomb's problem I will conclude that EDT cannot distinguish auspiciousness from efficacy.

One objection to this reasoning is, essentially, to ask: "if you're so smart, why ain'cha rich" (Lewis, 1981, p377)? Lewis's response to this is to argue that Newcomb's problem inherently rewards irrationality (Lewis, 1981, p377). Lewis argues for this on the basis that a Newcomb problem cannot be defined such that it rewards *U-rationality*, where *U* is "a kind of expected utility defined in terms of causal dependence as well as credence and value" (Lewis, 1981, p377). That is, a Newcomb problem cannot be defined such that it does not reward *V-rationality*, where *V* is "a kind of expected utility defined in entirely non-causal terms" (expected utility defined in terms of causal dependence as well as credence and value" (Lewis, 1981, p377). Lewis's argument takes the form of a *reductio ad absurdum* (Lewis, 1981, p379). Lewis's point is that, given the predictor's announcement that the *V-rational* choice is to be made *V-irrational* then the previously *V-irrational* choice becomes *V-rational*. So the predictor makes the same choice both *V-rational* and *V-irrational* (Lewis, 1981, p379). More formally, Lewis shows that upon the assumption that *V-rational* choice is to be made *V-irrational*, whether the agent construes the probability of getting the \$1,000,000, given that they take one box, as being greater than, equal to, or less than the probability of getting the \$1,000,000 if they choose both boxes, a contradiction always arises (Lewis, 1981, p379). Thus, Lewis has shown that the Newcomb problem does seem to inherently reward irrational behaviour.

CDT recommends two-boxing in Newcomb problems because in situations where the act has no propensity to affect the state, unconditional probabilities are used to calculate which action is best (Resnik, 1987, p113). So the utility of taking both boxes is $(\$1,000 \times p) + (\$1,001,000 \times (1-p))$ and the utility of taking just one is $(0 \times p) + (\$1,000,000 \times (1-p))$. Thus, CDT recommends two-boxing. EDT on the other hand falsely assumes that the probability of the outcomes is determined by the agent's choice so the expected utility of taking one box is $(0 \times 0.1) + (\$1,000,000 \times 0.9)^{15}$ and the expected utility of taking both is $(\$1,000 \times 0.9) + (\$1,001,000 \times 0.1)$, so EDT recommends one-boxing.

Eells's claim that EDT recommends two-boxing is based upon his 'tickle defence' of EDT, which rests upon his conception of an ideally rational agent. Eells assumes when they make their decision the agent knows what their subjective probabilities and desires are:

$$P(R\Phi(DM))=1 \text{ }^{16}$$

where 'RΦ(DM)' stands for the subjective probabilities and desires of the agent at the time of decision. He also assumes that because the agent is rational; the agent always performs the action he has determined to be correct:

$$P(BSA(DM) \leftrightarrow SA(DM)) = 1 \text{ }^{17}$$

where 'BSA(DM)' stands for the proposition that the agent has determined that SA(DM) is the correct act and 'SA(DM)' stands for the proposition that the agent performs the symptomatic act (SA). Thirdly, Eells argues that the agent's subjective beliefs and desires constitute a 'tickle' – an information screen – which prevents the normal probabilistic relationship operating between a common cause(CC) and the SA:

$$P(BSA(DM) | R\Phi(DM) \& CC(DM)) = P(BSA(DM) | R\Phi(DM) \& \neg CC(DM))^{18}$$

¹⁵ I have assumed that the probability of the predictor predicting accurately is 0.9.

¹⁶ see Eells 1982: 158

¹⁷ see Eells 1982: 159

¹⁸ see Eells 1982: 160

Eells argues for his third premise on the basis of the agent's rationality and a distinction between two different types of belief regarding the probabilistic relationship between CCs and SAs. These are called type-A and type-B beliefs. Type-A beliefs describe the probability relations between CCs and SAs for any random agent a (Eells, 1982, p154). Type-B beliefs describe such relations, but from an agent-relative perspective (Eells, 1982, p154). Eells claims that "type-B beliefs, and not type-A beliefs, should figure in the agent's calculation of the conditional expected utility of his available acts" (Eells, 1982, p154-155). This is because, given that the agent is rational, their course of action is dictated by their subjective probabilities, desires and a decision rule and the CC can only affect the agent's decision by altering one of these. Therefore, from the agent's perspective "there is no longer any motivation for an assumption of probabilistic relevance of the common cause to the determination of a given act to be the correct one" (Eells, 1982, p162).

From these three assumptions Eells is able to derive following two results:

$$\begin{aligned} P(\text{SA}(\text{DM}) \mid \text{CC}(\text{DM})) &= P(\text{SA}(\text{DM}) \mid \neg\text{CC}(\text{DM})) \\ P(\text{CC}(\text{DM}) \mid \text{SA}(\text{DM})) &= P(\text{CC}(\text{DM}) \mid \neg\text{SA}(\text{DM}))^{19} \end{aligned}$$

As Horwich points out, these final two results entail that there is no longer any evidential link between a CC and a SA. (Horwich, 1987, p186) This means the act of taking the extra \$1000 in Newcomb's problem does not imply anything about the predictor's decision (whether or not to place the \$1,000,000 dollars in the other box). Thus, the agent should take both boxes because the agent's subjective probabilities and desires have formed an information screen.

Eells's first two assumptions seem controversial to me. The assumptions that the agent knows $R\Phi(\text{DM})$ and that the agent will perform the act their considerations conclude is the best are standards that ordinary agents do not meet, at least not consistently. It is clear that ordinary agents do not necessarily know $R\Phi(\text{DM})$ at the time they make their decision. Eells second assumption seems to ignore the possibility of the trembling hand. Thus, Eells's theory only seems to apply to hyper-rational agents. However, Eells is not the first decision theorist to posit conditions of rationality that are not met by ordinary agents. Therefore, I shall

¹⁹ see Eells 1982: 162

focus my criticism upon Eells's third assumption, which, as Horwich shows, is more problematic.

Horwich's objection to Eells's third assumption is that $R\Phi(DM)$'s mediation of the causal chain between CC and SA does not necessarily give it any epistemological significance (Horwich, 1987, p187): that an agent's knowledge of $R\Phi(DM)$ conveys everything of some CC's causal relevance to his act does not entail he has knowledge of all the information that might be provided by a knowledge of the CC, or that his knowledge of $R\Phi(DM)$ encompasses all the evidentially relevant information to his eventual act that might be provided by knowledge of the CC (Sobel, 1994, p49). It might be objected to this that "an agent *should* be able to draw inferences from his discovery of... [$R\Phi(DM)$] about the probabilities of the alternative acts" (Horwich, 1987, p187). However, in order to determine the probability of an action relative to $R\Phi(DM)$, we must apply a decision rule using $R\Phi(DM)$ as our premises. The probability of various actions derived from this process cannot then be used in the deliberation about which act the agent will choose because the agent now knows which action he is going to perform (Horwich, 1987, p187). Furthermore, if the agent knows that he is not going to perform a certain action as a result of this deliberation, the expected utility of that action can no longer be calculated since it would "involve conditional probabilities relative to a condition *known* to be false" (Ibid). If no decision rule is applied to $R\Phi(DM)$, then nothing about probabilities of any of the agent's acts can be inferred and so on this reading Eells's third assumption is false (Ibid). Therefore, Eells's third assumption must be rejected, along with his conclusion.

In light of Horwich's objection to his third assumption Eells has presented a dynamic version of his tickle defense. This differs from his static tickle defense in that the agent does not simply assign probabilities to acts on the basis of deliberating upon $R\Phi(DM)$, rather this process becomes recursive: the agent reflects on $R\Phi(DM)$ which provides him with evidence of which act he would perform were these his final subjective probabilities and desires, "after conditionalizing on this evidence, he thinks again" (Sobel, 1994, p51). This process is repeated until $R\Phi(DM)$ has become stable. Thus, Eells's answer to Horwich's objection is to claim that a rational agent calculates the expected utility of an action through an alteration of the subjective probabilities contained in $R\Phi(DM)$ based on his previous calculations (it is in this sense the process is dynamic) (Sobel, 1994, p51). In the end an agent's "knowledge of his credences

and preferences and of what he is about to make of them screens off those causes from his acts, and makes them probabilistically independent of one another” (Sobel, 1994, p51).

Sobel’s main objection to Eells’s dynamic tickle defense of EDT is that it does not work as a general theory of rational decision because it does not give the correct answers to nondominance problems involving ordinary agents (Sobel, 1994, p52). First, Sobel considers the case of “Uncle Albert.” Albert is feeling unwell and has to decide whether or not to go to the doctor: on the one hand going to the doctor is a symptomatic act of being genuinely unwell but, on the other hand, going to the doctor will not make him ill if he is not already so, and staying at home will not improve his health if he is not well. Albert correctly decides to go to the doctor. However, an ideally rational agent, as stipulated by Eells, would decide to stay at home “and take this good news as evidence that he is going to stay at home and that he is not sick afterall!” (Sobel, 1994, p53).

Sobel assumes that the probability of Albert being genuinely ill is 9/15 and that acts of Go and ~Go are equally good further evidence as to whether or not Albert is Ill. News of either Go or ~Go make a 5/15 difference to the probability of the outcome for which it would be evidence. Given these assumptions ~Go would be the best choice based on the rational agent’s subjective probabilities, and so would be selected by Eells’s ideally rational agent, whilst Go would be the best act (Sobel, 1994, p64):²⁰

$$\text{Des}(\text{Go}) = -84 \text{ and } 2/3 < \text{Des}(\sim\text{Go}) = -31 \text{ and } 2/3$$

And,

$$\text{Util}(\text{Go}) = -58 > \text{Util}(\sim\text{Go}) = -60$$

It is clear that both Eells’s dynamic and static ‘tickle defenses’ of EDT fail. It is therefore extremely implausible to claim that EDT advocates a two-boxing strategy as a solution to Newcomb’s problem. We showed earlier that the rational strategy when faced with Newcomb’s problem is two-boxing. Therefore, CDT is a better model of rationality than EDT because it advocates two-boxing in Newcomb’s problem. This difference between the two ultimately results from a fundamental failing inherent to EDT; EDT is unable to distinguish

²⁰ See Tables 1 and 2 for further details.

auspiciousness from efficacy. Of course the evidential decision theorist can comfort themselves with their irrational gains, but these gains do not support EDT; they are simply proof that certain situations reward irrationality.

References

Eells, Ellery, 1982. Rational Decision and Causality. Cambridge: Cambridge University Press.

Horwich, Paul. 1987. Asymmetries in Time. Cambridge, MA: MIT Press.

Lewis, David. "Why ain'cha rich?", *Noûs*, Vol. 15, No. 3 (Sep., 1981), pp. 377-380

Resnik, Michael, 1987. Choices: An Introduction to Decision Theory. Minneapolis: University of Minnesota Press (2008 edition)

Sobel, Jordan, 1994. Taking chances: Essays on rational choice. Cambridge: Cambridge University Press.

Weirich, Paul, "Causal Decision Theory", *The Stanford Encyclopedia of Philosophy* (Winter 2008 Edition), Edward N. Zalta (ed.),
URL: <<http://plato.stanford.edu/archives/win2008/entries/decision-causal/>>.

DOES CRAIG'S KALAM ARGUMENT PROVIDE A CONVINCING ARGUMENT FOR GOD'S EXISTENCE?

JONANTHAN PUGH

The Kalam argument is a form of cosmological argument which argues from the hypothesis that the universe had a beginning in time, to the conclusion that God exists. In order to reach this conclusion, the argument relies on the crucial premise that everything which begins to exist has a cause. This paper shall argue that Craig's defence of the Kalam argument fails to offer a convincing proof of God's existence since it is possible to reasonably reject this causal principle.

The first section of this paper shall outline the premises of Craig's Kalam argument and qualify what is meant by a 'convincing proof', before appropriately limiting the scope of the issues to be considered. Craig outlines his version of the Kalam argument syllogistically as follows:

(1) Everything that begins to exist has a cause of its existence.

(2) The universe began to exist.

Therefore,

(C) The universe has a cause of its existence.²¹

Craig claims to have provided a convincing argument for God's existence since he claims that premises (1) and (2) have ample justification²² to compel us to accept their veracity. As such, for Craig, the argument "demonstrates the

²¹ Craig, "The Kalam Cosmological Argument", Wipf and Stock, (1979), p. 63

²² Ibid, p. 140

existence of a first cause which transcends and creates...finite reality".²³ Accordingly, this paper shall delineate the necessary and sufficient conditions of a 'convincing proof' as follows: A deductive proof can be deemed convincing iff it is valid, and its premises have 'ample justification' to compel us to accept their veracity. To clarify this second clause, we may say that one is compelled to accept that a premise is true if one is unable to reject it on the basis of either empirical evidence or tenable rational argument; such premises are true beyond reasonable doubt. Consequently, the Kalam Argument offers a convincing proof of God's existence for Craig; even those who initially doubt God's existence will be unable to deny the argument's conclusion, unless they abandon a premise which we seem compelled to accept.

Before beginning the argument of this paper, it is important to highlight some limitations that this essay shall make. It seems that one could raise a Humean objection²⁴ to the Kalam argument insofar as it does not explicitly prove the God of monotheism as opposed to polytheism, or that God must have divine attributes. This Humean protest is a fair objection; however, due to limitations of space it shall be ignored for the purposes of this essay.²⁵ To clarify, this paper is investigating the proof of a creative force which brought the universe into existence, not a single deity with divine attributes.²⁶

Moreover, this paper shall grant Craig his second premise (that the universe is finite). This shall be done for two reasons; firstly, although there are potent objections to Craig's a priori arguments against the possibility of an infinite temporal sequence²⁷, these objections do not render his second premise untenable. We may instead accept premise (2) because of the empirical evidence for a finite universe which Craig also provides in his exposition.²⁸ Therefore, it shall be assumed that the universe did have a beginning.²⁹ However, it is important to explain a key concept of Craig's argument here. By Craig's

²³ Ibid, p. 64

²⁴ This objection is aimed at teleological arguments in Hume, "Dialogues Concerning Natural Religion", Nelson & Sons, (1947), p.166

²⁵ For Craig's response see Craig (1979) pp. 172-174

²⁶ Craig offers arguments for the claim that the cause of the universe must be personal; see Craig (1979) pp (149-152). This does not affect this paper's argument.

²⁷ See Smith "Infinity and the Past", *Philosophy of Science*, 54, (1987), pp. 63-75

²⁸ Craig (1979), pp 110-140

²⁹ However, this acceptance is not unanimous; see Hawking, "A Brief History of Time", Bantam Press, (1988), pp. 136-141

arguments, a crucial distinction can be drawn between the universe and God; for Craig, the universe has to have had a beginning, whilst God is timeless.³⁰ This is pertinent to the argument being discussed here, since it allows Craig to claim that whilst the universe requires a cause, God does not; this is because only things that begin in time require a cause of their existence.

Having made these preliminary remarks, the main argument of the paper can begin. The overall aim of this paper is to argue that Craig fails to offer a convincing proof of God's existence because we are not compelled to accept premise (1), (referred to henceforth as 'the causal principle'). The aim of the second section of this paper is to provide a comprehensive exposition of Craig's arguments for the causal principle. Craig's foremost argument here is that the causal principle is "intuitively obvious"³¹; however, he supplements this argument by indicating two further ways in which one could elaborately defend the principle. In the interests of thoroughness, this section shall begin by showing that the supplementary arguments which Craig indicates do not demonstrate the truth of the causal principle. Having exposed the flaws of these arguments, it shall be claimed that Craig's assertion of the causal principle relies on his main argument from metaphysical intuition, which presents a stronger argument than his supplementary claims.

Craig's first supplementary argument claims that the causal principle can be supported by empirical facts. Since nobody has encountered an effect which was not preceded by some cause, and every event we have experienced has been caused, the causal principle has overwhelming empirical support.³² However, as Craig himself notes, this argument is philosophically weak;³³ a Humean sceptic may respond that no matter how many instances we may observe of causal relations holding, it does not entail that this relation holds necessarily. It is logically conceivable to imagine an effect which exists without a cause.³⁴ Therefore, since there is no necessary connection between a cause and an effect, it is not logically impossible that the causal principle does not apply in

³⁰Ibid, p. 170 ft 159. See also, Craig, "God, Time and Eternity", Dordrecht, Kluwer, (2001)

³¹ Craig, (1979), p. 141

³² Ibid, p.145

³³ Ibid, p. 145

³⁴ See Hume, "An Enquiry Concerning Human Understanding", Oxford University Press, ed. Millican (2007), sections IV and VII

environments which we have not experienced, such as the initial state of the universe.

However, this is not an unassailable objection to Craig's use of the causal principle. If we rejected empirical generalisations on the basis that they are not analytically true, we would have to be sceptics about most forms of synthetic knowledge. However, such scepticism does not seem to be a tenable position; in everyday life we accept synthetic claims such as the causal principle, because we believe that the empirical evidence in their favour has demonstrated their truth beyond reasonable doubt. As such, one could argue that we must agree that we are compelled to accept the causal principle by empirical evidence, if we are to avoid an untenably sceptical position.

Despite this defence, Craig's first supplementary argument is still flawed. The flaw lies in the fact that the evidence of the causal principle applying in the universe does not give us any relevant evidence for the hypothesis that it applies to the universe itself. Indeed, to use this evidence to establish that the causal principle applies to the universe itself is to commit the fallacy of composition,³⁵ whilst we may generalise from the evidence of all effects in the universe being preceded by causes, to the conclusion that causal laws will probably continue to apply in the universe, we cannot use this evidence to validly derive that the universe itself was caused.

Craig's second supplementary argument is an endorsement of Hackett's neo-Kantian claim that the mental category of causation necessarily represents reality. In Kantian metaphysics, reality consists of two separate realms; the phenomenal world, and the unknowable 'noumenal' world of 'things in themselves'. According to Kant, we obtain knowledge of the phenomenal world through sense data, and the a priori categorical structure of the mind. These 'categories' of the mind are objectively valid epistemic conditions; if we did not perceive the phenomenal world with these categories in place, we could not come to have knowledge of things as they appear to us. Therefore, the assumption of the mental category of causality is essential if we are to apprehend, and have knowledge of the phenomenal world.³⁶

³⁵ Russell makes a similar claim; see transcript of Russell/Copleston debate at http://www.scandalon.co.uk/philosophy/cosmological_radio.htm (1948)

³⁶ This paragraph summarizes Craig, (1979), p.146; (paraphrasing Kant, Critique of Pure Reason, cited below)

Hackett goes beyond these Kantian claims and asserts that the category of causality must also apply to the noumenal realm of things-in-themselves. His first argument for this is that the Kantian position of the categories applying only to the phenomenal world is “self refuting.”³⁷ If we follow Kant in accepting that the operation of the categories is limited to the phenomenal world, then it would be impossible to have knowledge of the categories, since they are “characterised by the absence of sense data.”³⁸ However, since (according to Hackett) we do have knowledge of the categories, they must therefore apply in the noumenal world of things in themselves.³⁹ Secondly, Hackett argues that since the noumenal thing-in-itself causes the appearance of perceptual objects in the phenomenal realm, the category of causality must consequently apply between the noumenal and phenomenal realms at the very least.⁴⁰ Therefore, for Hackett, Kant’s metaphysics can be extended in a way which demonstrates that causality must be a feature of reality in itself.

However, Hackett’s arguments fail to establish the truth of the causal principle because they conflate two separate aspects of Kantian metaphysics. We can identify three different planes in Kant’s metaphysics; we may classify the noumenal realm as the ‘Transcendent’ plane, and the phenomenal realm as the ‘Empirical’ plane. The Kantian categories can be classified as belonging to a separate plane, which we may term the ‘Transcendental’. The ‘Transcendent’, and the ‘Empirical’ planes are constitutive of reality; that is, for Kant, reality consists of just the phenomenal and noumenal realms. In contrast, the ‘Transcendental’ plane is regulative; the concepts it postulates do not constitute reality, but rather regulate how we come to apprehend reality.⁴¹

With this terminology in mind, it seems that Hackett conflates the regulative ‘Transcendental’ plane of Kant’s metaphysics with the constitutive planes. This misunderstanding pertains to the epistemic status of the categories. For Kant, the

³⁷ Craig (1979), p.146, (paraphrasing Hackett, “The Resurrection of Theism”, Chicago: Moody Press, (1957)

³⁸ Ibid, p. 146 (paraphrasing Hackett)

³⁹ Ibid, p. 146 (paraphrasing Hackett)

⁴⁰ Ibid, p. 147.

⁴¹ I thank Dr. Dominic Smith for clarifying this terminology. See also Allison “Transcendent Idealism”, Yale University Press, (1983).

categories are never applied to the 'Transcendent' noumenal plane, but are "always only of empirical use"⁴² as such, we cannot know that they apply to the transcendent things-in-themselves. However, Hackett makes the unwarranted claim that we actually know that the categories apply in reality⁴³. According to Kant, this cannot be the case. We must act as if the categories obtain between the 'Transcendent' and 'Empirical' planes; however, this regulative relationship is merely a "postulation."⁴⁴ The relationship is not constitutive of reality, and is therefore not something that can be known.⁴⁵

Accordingly, contra Hackett's first argument, Kant does not contradict himself in claiming that the causal category is limited in its operation to the phenomenal world; we just do not have knowledge of the categories as Hackett assumes. Moreover, Hackett's second argument is flawed because it implicitly assumes that things in the noumenal realm cause the appearance of objects in the phenomenal realm in a constitutive sense. However, for Kant, this is not the case; rather, the noumenal realm merely provides the necessary and sufficient regulative conditions for us to apprehend perceptual objects. Hackett fails to provide a valid reason to assume (contra Kant) that this causal relation is constitutive, rather than merely regulative of reality. As such, although the category of causation might be a necessary feature of our mental representation of reality, Hackett's arguments do not demonstrate that causation is a necessary feature of reality itself.

Therefore, neither of Craig's supplementary arguments demonstrate that the causal principle is indisputable; we cannot be certain that it applies to the universe itself on the basis of these arguments. Therefore, the final part of this section shall explicate Craig's main case for the causal principle, and explain why it presents a stronger argument in its favour.

For Craig, the real force behind the assertion of the causal principle is that it is a metaphysical intuition; although there is no logical contradiction in postulating an effect without some sort of cause, to claim that this could actually occur in the real world is to agree to something "worse than magic"⁴⁶ The thought that the

⁴² Kant, *Critique of Pure Reason*, ed. Guyer, Cambridge University Press, (1998), p. 345 (*B303*).

⁴³ See footnote 19.

⁴⁴ Allison, (1983), p.4.

⁴⁵ Kant, (1998), see Book 2, Ch 3 for Kant's account of this argument.

⁴⁶ Craig, "Blackwell Companion to Natural Theology", Wiley-Blackwell, (2009), p. 186.

universe was caused to exist is just “eminently more reasonable”⁴⁷ than its springing into existence uncaused. Therefore, for Craig, we are compelled to accept the Causal Principle because of our “[...] metaphysical intuition that something cannot come out of nothing.”⁴⁸

According to this argument, the fact that the causal principle is not logically necessary is not problematic. One can concede that the denial of the causal principle does not involve a logical contradiction, yet maintain that its rejection would involve assenting to something which is intuitively absurd. As Craig asserts, “there are other absurdities than logical ones.”⁴⁹ Moreover, this argument seems to avoid the second objection to the first supplementary argument. This objection claimed that one commits the fallacy of composition if one argues from the fact that every proper part of the universe is caused, to the conclusion that the whole universe is caused. However, it does not seem that Craig commits this fallacy in his main argument. For Craig, the empirical support of the causal principle gives us further evidence to suppose that it is true; however, this is not the reason that we actually adopt the principle. Instead, we adopt it because of its inherent intuitive metaphysical plausibility. Therefore, although Russell is correct to highlight that it is logically invalid to derive the application of the causal principle to the universe from its application to things in the universe, Craig may respond that this inference does not figure in his main argument for the causal principle.

To take stock, the second section of this paper has demonstrated that although Craig’s supplementary arguments for the causal principle are flawed, this is not in itself a telling objection to his use of the causal principle. The crux of Craig’s argument is that we adopt the causal principle because of its intuitive plausibility, aside from the questionable supplementary arguments in its favour. Therefore, in responding to Craig, we must object to his use of the causal principle when it is assumed as a highly intuitive metaphysical principle. The final section of this paper shall begin by describing two possible lines of response to Craig’s main argument here. It shall ultimately be concluded that the reason that his proof fails to be convincing is that Craig fails to establish that the

⁴⁷ Ibid, p. 144.

⁴⁸ Craig, “The Caused Beginning of the Universe”, *British Journal for the Philosophy of Science*, 44:4, (1993), p. 628.

⁴⁹ Craig, (1979), p. 145.

uncaused origin of the universe must be less intuitively plausible than the theistic alternative.

In contemplating the origin of the finitely old universe, it seems that we must choose between two mutually exclusive sets of beliefs. On one hand, we may believe that the causal principle applies to the finite universe, which thereby commits us to belief in a supernatural creator of the universe. Alternatively, we can avoid this theistic conclusion, but then be committed to the rejection of the causal principle. The thrust of Craig's argument is that the causal principle is so intuitively obvious, that we are compelled to accept the first set of beliefs over the second. Accordingly, it seems that there are two strategies which one can adopt in order to argue that we need not be compelled to accept the causal principle. Firstly, one could argue that something coming into existence uncaused is more plausible than Craig describes. Conversely, one could argue that the plausibility of the caused beginning of the universe is less plausible than Craig describes. The first part of this section will give an account of recent attempts to carry out the first strategy, before explaining why it fails. However, it shall be concluded that the second strategy provides a powerful line of objection to Craig's adoption of the causal principle, and accordingly shows why his Kalam argument is not a convincing proof of God's existence.

It seems that quantum mechanics has provided evidence which increases the plausibility of something having an uncaused existence. According to Heisenberg's Uncertainty Principle, the nature of quantum systems implies that the precision with which one can determine the energy of a particle is inversely related to the precision with which one can determine the time at which the particle possesses that energy. As such, it is impossible to calculate a precisely determined value for both the amount of energy a particle has, and the duration for which it has this energy.⁵⁰ This is pertinent for the following reason; if the duration of time that a particle possessed some energy is measured with a high degree of certainty, then by Heisenberg's uncertainty relation, the degree of uncertainty concerning the energy of the particle over that time period will be so large, that:

⁵⁰See Hilgevoord, "The Uncertainty Principle", at <http://plato.stanford.edu/entries/qt-uncertainty/> (2006)

"[...] it becomes impossible in principle to determine if the law of conservation is violated...and consequently an amount of energy can spontaneously come into existence."⁵¹

Moreover, there is indirect observational evidence⁵² that the spontaneous creation of these 'virtual particles' "frequently occurs."⁵³ As such, it has been asserted that "[...] energy can appear out of nowhere"⁵⁴ spontaneously, which seems to imply that there are counter examples to the causal principle; virtual particles can emerge from quantum vacuums without an apparent cause. Furthermore, one could attribute the creation of elementary matter to these fluctuations of energy in quantum vacuums⁵⁵. Thus, it seems that this evidence might provide us with a reason to reject the causal principle, and its application to the origin of the universe.

However, these purported counter examples to Craig's premise fail to acknowledge an important aspect of his use of the causal principle. Virtual particles provide examples of existents which lack sufficient conditions; that is, we can list all the conditions which were met when a virtual particle is said to have existed, without being able to determine from that information alone whether or not the particle will exist if those conditions are replicated. However, the thrust of Craig's Kalam argument lies in the intuitive nature of the notion that there must be necessary conditions for things which come to exist. Crucially, virtual particles are not examples of existents which lack necessary conditions⁵⁶, in that they still require the existence of quantum energy from which to arise; they do not strictly arise from 'nowhere' as the Davies quote intimates.

Therefore, Craig can reply that there is a crucial difference between the origin of a virtual particle and the origin of the universe. As Craig explains, "the quantum vacuum is not nothing;"⁵⁷ although virtual particles seem to spring into existence from 'nowhere', in reality they are the result of the undetermined fluctuations of

⁵¹ Smith, "The Uncaused Beginning of the Universe", *Philosophy of Science*, 55, (1988), p.50

⁵² See Lamoreaux, "Demonstration of the Casimir Force in the 0.6 to 6 μ m Range", *Physical Review*, 78:1, (1997), pp.5-8

⁵³ Smith, (1988), p. 50

⁵⁴ Davies, "God and the New Physics", J.M. Dent & Sons (1983), p. 162

⁵⁵ See Guth, "The Inflationary Universe", Basic Books, (1998)

⁵⁶ Craig (1993), pp. 627-628

⁵⁷ Craig (2007), p. 1

a pre-existing energy. In contrast, it seems that the universe must have originated independently of any pre-existing matter or energy. As such, it seems that even virtual particles require a cause in the sense that Craig believes is relevant, as this quotation makes clear:

“If nothing existed —no matter, no energy, no space, no time, no deity—[...] then it seems unintelligible to say that something should spring into existence.⁵⁸

Furthermore, it implicitly follows from Craig’s argument here that it is impossible for us to ever gather empirical evidence of truly uncaused events. The reason for this is that it is impossible to replicate a perfect vacuum in which absolutely nothing exists. The only empirical evidence which would refute Craig’s sense of the causal principle is if something was observed spontaneously coming into existence in such a vacuum; yet this perfect vacuum is only a conceptual possibility within the universe. Therefore, it seems that we can only refute Craig’s causal principle theoretically, and not empirically.

Consequently, empirical attempts to increase the plausibility of an uncaused universe fail because they cannot provide examples of existents which lack necessary conditions for their existence. Moreover, any attempt to increase the plausibility of uncaused existences through purely theoretical physics would surely meet strong opposition from Craig; as he intimated in his discussion of actual infinites when defending premise (2), one can accept that something is mathematically possible, but argue that it cannot occur in reality⁵⁹. Anticipating this potential response for Craig, the remainder of this paper shall instead adopt the second strategy outlined at the beginning of this section in attacking Craig’s argument. In doing so, it shall aim to provide an objection to Craig’s Kalam argument from a different perspective, by arguing that we are not compelled to accept the causal principle, because the supernatural implications of accepting that the universe had a cause are far more implausible than Craig describes.

Before expounding this argument, it is essential to dispel an anticipatory defence that Craig provides against such an objection. Craig writes that “only an aversion

⁵⁸Craig (1993), p. 627

⁵⁹ Craig (1979), p. 72

to the theism implied by the principle⁶⁰ could lead one to reject the causal principle. This seems to insinuate that the objector to the causal principle is invalidly clinging to the notion of an uncaused universe because he does not want to believe in a supernatural creator. This rejection of the causal principle would of course be unfounded; we are not justified in rejecting a clear truth just because of an inclination not to believe its implications. However, it seems that one can be averted to a conclusion in a manner which would allow one to validly reject it. Specifically, one may be averted to a conclusion because of its inherent implausibility; that is, one may not rationally be able to bring oneself to believe such a conclusion, despite the evidence in its favour. Thus, whilst rejecting the causal principle because of one's disinclination to theism is invalid, it seems that one can validly reject the principle if one can provide rational grounds for the rejection of theism. Therefore, the objection must reject the causal principle because of a rationally justified aversion to theism, and not just on the basis of a disinclination towards theism.

It seems that such a rational justification can be provided. As was explained in the first section of this essay, the God which the Kalam argument posits must transcend time; only then can we make sense of God not requiring a cause. Such a view of God may be coherent to the theist⁶¹; however, it is far from being an intuitive concept. Such a being seems shrouded in mystery; we might ask 'how can something exist outside of time?' and 'how can such a being create a temporal universe?'. Craig himself admits that a timeless God is a "mind-boggling conclusion⁶²", yet claims that it is more plausible as an explanation for the origin of the universe than an uncaused universe. However, one can surely offer the rejoinder that the notion of an uncaused universe might just be more credible than a highly implausible transcendent creator to a non-believer.

A similar thought is implicit in Mackie's objection to the Kalam argument. Mackie writes that the Kalam argument just assumes

"...that God's existence and creative power (is) self explanatory, whereas the unexplained origination of a material world (is) unintelligible."⁶³

⁶⁰ Ibid, p.628, ft 5.

⁶¹ See Craig (2001).

⁶² Craig (2007).

⁶³ Mackie, J. "The Miracle of Theism", Clarendon Press, (1982), p. 94.

Unfortunately, the terminology Mackie uses here is inappropriate. For Craig, it is not the case that God must be self-explanatory; all that Craig need claim is that God does not require a cause, since only things which begin to exist in time require a cause. As such, the criticism Mackie goes on to make is misguided; according to Mackie, the concept of a self explanatory God is “not defensible⁶⁴” since it relies on the questionable reasoning of the ontological argument⁶⁵. Yet, as has been explained, Craig is not committed to this conception of a self explanatory God⁶⁶. However, although the semantics of Mackie’s objection are flawed, the spirit in which it is made is not. It seems that we could grant to Craig that an uncaused God is logically coherent, yet attack the assumption that the concept of an uncaused God is intuitively intelligible, whilst an uncaused universe is not. To a non-believer precisely the opposite may have far more intuitive appeal.

Craig claims that we are compelled to accept the causal principle because of its intuitive force. The strength of the objection being made here is that it uses Craig’s appeal to intuition against his own argument; although one may have an intuition that causation applies to the universe, it is perfectly conceivable that one may have an equally strong intuition that a transcendent being cannot exist; in fact, the implausibility of the deity’s timeless nature provides reasons which can justify why one may have this anti-theistic intuition. For theists, the incomprehensible nature of God need not be problematic, since God is accepted as being beyond human comprehension. However, this idea of God is problematic with respect to the strength of Craig’s proof, since the argument relies on the thought that the causal principle is the most intuitive hypothesis to adopt in explaining the origin of the universe. Yet, if one maintains that the idea of God is so contrary to one’s reason as to make the rejection of theism preferable to the preservation of the causal principle, it seems that one can have rational grounds to reject the causal principle.

Admittedly, the scope of this objection is somewhat narrow; it will not deplete the strength of Craig’s argument for those who believe that the concept of God is

⁶⁴ Ibid, p. 94

⁶⁵ Ibid, p. 84

⁶⁶ A self-explanatory God need only be the conclusion of an argument based on the Principle of Sufficient Reason, not the Causal Principle. For such an argument see Leibniz “On the Ultimate Origination of Things” in Leibniz: Philosophical Writings, ed. Parkinson, Guernsey Press, (1973), pp. 136-145.

not counter-intuitive. However, since there are reasons that can justify why one may find such a being counterintuitive, the objection shows that the Kalam argument fails to be a convincing proof of God's existence, on this paper's definition. Craig's argument turns on the thought that we are compelled to accept the causal principle because of its intuitive strength; however, this reliance on intuition leaves it open to the non-believer to counter that he does not, and more importantly need not have the same intuition.

The theist might point out that those who are not convinced by Craig's Kalam argument have failed to appreciate just how counter intuitive the rejection of the causal principle is; no matter how mysterious God may seem, he can surely not be more implausible than the rejection of the causal principle. However, this reply just begs the question. It assumes that the rejection of the causal principle must be more implausible than the acceptance of theism; yet this is exactly what needs to be demonstrated. Moreover, the theist cannot appeal to the evidence of empirical generalisation to establish the truth of the causal principle without committing the fallacy of composition as was argued earlier. As such, if the only argument we have for the causal principle is the strength of our intuition, then it seems that this argument can be defeated by a stronger anti theistic intuition. Therefore, if the Kalam is to provide a convincing proof of God's existence, it must demonstrate that the causal principle and its theistic corollary is the only intuition that a fully rational person can have concerning the origin of the finite universe. It just does not seem that Craig's arguments succeed in this task.

Therefore, Craig's Kalam argument fails to provide a convincing proof of God's existence. Although we may lack empirical evidence which demonstrates that the causal principle is not universally applicable, it still seems that one can coherently reject it given its implausible implications. The rejection of the causal principle is only valid if it is not based on a mere disinclination to theism; this paper has argued that since the reason for accepting the causal principle is its intuitive force, it can be coherently rejected if one does not accept that the conjunction of the causal principle and its theistic implication has a greater intuitive appeal than its rejection. Therefore, on this essay's stipulated conditions of a 'convincing proof', the Kalam argument fails to provide a convincing proof of God's existence, since we are not compelled to accept the causal principle. As a closing thought, we may instead regard the argument as a way to expose where the intuitions of theists and non-believers concerning the origin of the universe diverge.

REFERENCES

Allison "Transcendent Idealism", Yale University Press, (1983)

Craig "The Kalam Cosmological Argument", Wipf and Stock, (1979)

Craig "The Caused Beginning of the Universe", British Journal for the Philosophy of Science, 44:4, (1993)

Craig "God, Time and Eternity", Dordrecht, Kluwer, (2001)

Craig "Blackwell Companion to Natural Theology", Wiley-Blackwell, (2009)

—<http://www.reasonablefaith.org/site/News2?page=NewsArticle&id=7279>, (2007)

—<http://www.reasonablefaith.org/site/News2?page=NewsArticle&id=5673>, (2007)

Davies "God and the New Physics", J.M. Dent & Sons (1983)

Guth "The Inflationary Universe", Basic Books, (1998)

Hackett "The Resurrection of Theism", Chicago: Moody Press, (1957)

Hawking "A Brief History of Time", Bantam Press, (1988)

Hilgevoord "The Uncertainty Principle", at
<http://plato.stanford.edu/entries/qt-uncertainty/>, (2006)

Hume "Dialogues Concerning Natural Religion", Nelson & Sons, (1947)

Hume "An Enquiry Concerning Human Understanding", Oxford University Press, ed. Millican, (2007)

Kant *Critique of Pure Reason*, ed. Guyer, Cambridge University Press, (1998)

Lamoreaux "Demonstration of the Casimir Force in the 0.6 to 6 μ m Range", Physical Review, 78:1, (1997)

Leibniz *Philosophical Writings*, ed. Parkinson, Guernsey Press, (1973)

Mackie, J. "The Miracle of Theism", Clarendon Press, (1982)

Russell, *Russell vs Copleston Radio Debate* transcript at
http://www.scandalon.co.uk/philosophy/cosmological_radio.htm (1948)

Smith "Infinity and the Past", *Philosophy of Science*, 54, (1987)

Smith "The Uncaused Beginning of the Universe", *Philosophy of Science*, 55,
(1988)

EDITORIAL AFTERWORD

Though this first edition is small, we feel it was important to get the ball rolling. To create a space within the philosophical community of the University of Edinburgh where primarily (but not restricted to) undergraduate students could submit their own work and have it assessed within a journal format. In this edition we have presented two papers in epistemology and decision theory, in the hope of stimulating more submissions from any of a broad range of topics within philosophy. Whilst it is hard to delineate exactly where the edges of philosophy lie, we should hope to encourage submissions not simply within the standard range (and we hope this list is not controversial) of moral philosophy, epistemology, meta-ethics, political philosophy, philosophy of mind, logic, aesthetics and epistemology, but also the philosophies 'of', such as philosophy of science or religion (though if a paper can be written presented a philosophy 'of' something more exotic we would encourage this also). If you have an idea which can be formulated into a philosophy argument of a standard you are proud of, please submit it to uepsj.editorial@gmail.com . Likewise we would welcome the submission of any work previously submitted as part of academic assessment. Simply put, if its good prose, outlining interesting philosophy and is around 3,000 words we want to see it.

The exact format of this journal also requires explanation. Though we have only two submissions at the time of writing, we none the less hope to expand upon that submission by submission as the time goes on. This 'rolling' submissions system does not mean that at times papers will be released in more traditional composite 'editions' (as circumstances dictate, for example if 6 strong political philosophy papers were accepted within a short time period, this might warrant an edition dedicated to political philosophy). However, if a paper is submitted, processed and accepted, we work upon the assumption it is better published than not. As such, we have at least partly adopted the 'rolling' style of content distribution. On this note we would also encourage the submission of reviews and replies to any papers within this journal and any future releases in the interest of promoting interesting dialectic (this being the method by which philosophical ideas progress). Thus, for this journal there will not be 'submission dates' strictly speaking. Instead (unless stated otherwise) we will be aiming for the time period from submission to publishing of a successful article to be dictated by the work load upon the journal staff alone.

Thirdly we would encourage students to be more open in regards to submitting work than they would perhaps be within normal academia. This is not to say there is anything like a anything goes mentality. Yet if you have an idea or piece that you might perhaps consider too controversial or fringe to submit for more formal assessment. Again, please consider this journal.

Finally, we would like to close this first edition with a challenge. This is not contrary to the above, rather we hope that be presenting broad topics as a challenge we might stimulate interesting discussion on the topic. As such, this issue we would like to present the question: *Why be moral?* In keeping with the themes expressed above we encourage the interpretation and content of any response to this question to simply be whatever makes the best argument. Is Kantianism the last word? Does Thrasymachus have a point? Is there a convincing argument for moral nihilism? Or any other way of approaching the question in a philosophical manner we have not thought of.

We eagerly await any work you are willing to submit.

Richard S. Gow
Chris Ranalli
Editors