

Undergraduate Philosophy

Good Answers Guide

UWA Undergraduate
Philosophy Society

2021

Introduction

The news is not new that good answers guides are rarer at universities than they are at high schools. In light of this, the UWA Undergraduate Philosophy Society is pleased to present this collection of outstanding undergraduate essays.

The guide is comprised of seven essays that achieved a grade of 80% or higher. We have organised these by unit level and unit code, see below:

Level One

- 3 PHIL1003 God, Mind and Knowledge: An essay that critiques Anselm's ontological argument.
- 9 PHIL1003 God, Mind and Knowledge: An essay that defends Cartesian dualism.

Level Two

- 15 PHIL2001 Bioethics: An essay that defends compulsory vaccination.
- 22 PHIL2006 Philosophy of Psychology and Psychiatry: An essay that critiques the classical computational theory of mind.

Level Three

- 29 PHIL3003 Moral Theory: An essay that defends hedonic act-consequentialism.
- 37 PHIL3003 Moral Theory: An essay that critiques the hedonism theory of wellbeing.
- 41 PHIL3005 Continental Philosophy: An essay that critiques Martin Heidegger's 'The Thing'.

Our club would like to give thanks to all contributors. Thank you to the authors who generously submitted their undergraduate assignments. Thank you also to the philosophy discipline that approved the project and provided the first essay. And, thank you to Brindy Donovan who designed and formatted the guide.

Jack Stewart
President 2021

Level One

PHIL1003: God, Mind and Knowledge

Anonymous

Essay Question: What is the ontological argument for the existence of God? Explain clearly how the argument is supposed to work. Does it succeed? If it does not succeed, explain what is wrong with it. If you think it does succeed, consider a challenge to it by one of the philosophers mentioned on this unit outline, and offer a defense against that challenge.

My thesis is that Kant's objection to Anselm's ontological argument (AO), that existence is not a real predicate, shows that AO does not succeed in trying to prove the existence of God. I shall argue for my thesis firstly by outlining and explaining AO. I shall then proceed to explain, in detail, Kant's famous objection. Next, I will discuss whether Kant's objection works, and look at a criticism (offered by Peter Millican amongst others) that his objection is irrelevant. Finally, I shall conclude in my thesis that AO does not succeed.

A good introduction will contain a clear and unambiguous statement of your thesis. The author states her thesis in the first sentence. It's okay to state your thesis later in the introductory paragraph. But this is not a bad way to go.

Anselm's argument, just like all other ontological arguments, is supposedly an a priori proof of God's existence – he aims to establish God's existence from the very concept of God. Anselm (1989:96) then goes further than this, and he argues that God exists in such a way that his non-existence cannot be conceived.

The author uses her introductory paragraph to give the reader a very concise overview of what to expect. (That's good.)

Notice also that the introduction doesn't contain any grandiose detours like 'Since the dawn of time, mankind has wondered about whether God exists.' It's just straight to business. This is the right sort of approach in a short philosophy paper. (Frankly, it's a good approach for all philosophy papers).

Anselm (1989:96) begins his argument by considering a non-believer which he calls the Fool. He says that this Fool understands Anselm's conception of God – something than which nothing greater can be thought. Anselm's reasoning can be shown clearly by these six steps:

Right. But there is more that can be said here on behalf of the theist. I have in mind a detail that makes it hard for the atheist to resist this particular claim. In particular, the atheist herself asserts the sentence 'God does not exist'. Anselm points out that, unless the atheist wants to admit that she doesn't understand what she herself is saying, then she has to concede that she has an idea of God—that is, God at least exists in her understanding.

- (1) The Fool understands Anselm's definition of God, "something, than which nothing greater can be thought" (AnselmGod).

- (2) Thus, AnselmGod exists, at least in the Fool's mind.
- (3) It is greater to exist in reality than to only exist in the mind.
- (C1) So, if AnselmGod existed only in the Fool's mind, then it would be possible to think of something greater (for example, the same thing existing in reality as well as in the Fool's mind).
- (5) But this would be a contradiction, since by definition, it is not possible to think of something greater than AnselmGod.
- (C2) Therefore AnselmGod must exist both in the Fool's mind and in reality.

Nice recap of the argument.

Generally, I would like to see you follow this up with a little bit of commentary fleshing out the key premises and offering some reasons why someone would be tempted to believe them. (E.g., with respect to P3, you might offer an example that seems to support it: a real ice cream cone is better than a non-existing ice cream cone..)

What I find so fascinating about Anselm's argument is that it appears, *prima facie*, to be logically sound. It seems that even the non-believer must initially follow the argument, as nothing stands out to be a cause for concern immediately after reading the premises.

Here the author offers a "signpost" for the reader. She gives a quick, one-sentence description of what the next few paragraphs are about and/or how they fit into her discussion. This makes it much easier for the reader to follow the discussion.

I am now going to discuss what I take to be the strongest objection to AO, which comes from Immanuel Kant. The ontological argument rests on the assumption that a God who exists is greater than a God who does not exist. Kant (1989) argued that this rests on a confusion, as existence is not a real predicate. To clarify, a predicate is a property that a thing can either have or not have. Kant's line of thought goes as follows. When those who believe in God say "God exists", they are not saying this as a shorter way of saying "there is a God who has the property of existence". If that was what they actually meant, then non-believers such as the Fool would be saying, when they say that "God does not exist", that "there is a God who does not have the property of existence". This would be a contradiction due to both asserting and rejecting God's existence in the same sentence (Kant 1989: 106-109).

I had some misgivings about this claim right here. Predicates are a kind of word, whereas properties aren't words. They are, rather, qualities or attributes that objects have. That said, Kant himself conflates predicates and properties, and the student may here just be relaying this fact. But I wish she would have made it more explicit that she was just giving to Kant his mistaken usage. She could have written something like, "When Kant says existence is not a real predicate, what he seems to mean is that it is not a real property..."

(As you will see in a moment, the author does eventually make clear what is going on. But I think she should do it right here. Other wise, there is a risk that the reader can become confused.)

Instead, Kant (1989:107) suggests that when we say something exists, we are actually saying that the concept of the thing that we are talking about is displayed in the world. So, existence is

not a property that the thing can possess. Instead, it is a concept that corresponds to something in the world. To clarify, Kant uses predicate to mean property; I will use the two interchangeably throughout this essay.

I will now illustrate Kant's objection. It is important to first clarify what a property is. A real property is something which, when it is added to the complete concept of a thing, changes that concept. It will be clearer by explaining with an example – say, the concept of an orange ball. The complete concept of this orange ball contain these properties:

- Is orange
- Is spherical

Kant's point is that if we add 'exists' to the concept of the orange ball, nothing about that concept changes so existence cannot be a real property. Whereas if we were to add a property such as 'is shiny', the concept of the orange ball would change.

So Kant (1989) argues that we cannot prove any kind of existence of anything from its concept because existence is not a real predicate; adding existence to a concept does not define a new concept. There is no difference between the concept of something possible and the concept of something actual. Therefore, the ontological argument is not sound due to this mistaken assumption.

One response I shall now consider is that Kant's objection is irrelevant to Anselm's ontological argument. Granted, Kant's objection has been accepted by logic – existence is not logically a predicate. This is because when you represent an object and its predicate in logic, you would use the form 'there is some x which is P' as opposed to 'the thing which is A (being/exists) is also P'. Thus, according to the form of the former logical sentence, it appears that existence is assumed to not be a predicate. But even so, it does not mean it is relevant to AO (Millican 2004:437-438). Due to the scope of this essay, I shall now cover which premise of

Here the author remembers to clarify Kant's (mis)use of the word 'predicate'. But see my above comment.

More signposting. Good. In addition, notice that the author is now going to go into more depth in explaining Kant's argument. A lot of students just give a quick summary of an argument, and then move on to other things. In general, we like to see a paper in which a student explain fewer ideas in greater depth, rather than explain many ideas but on a surface level.

Right. It would be also good to note Kant's argument for this point: if 'existence' was just another component in a concept of a thing, then whenever we say that 'x exists' after earlier saying 'x does not exist' we wouldn't count as having changed our mind about a single thing x. This is because the referent/meaning of 'x' changes once we think of 'x' as picking out a thing that has the quality of existence.

My biggest criticism of the paper is that I never feel like it is clearly explained what the theist's objection is. Even as I look below, I'm not sure WHY the theist might say that Kant's objection is irrelevant. Why does the theist say it does not matter whether existence is a property? If this is Millican's argument you are making reference to, you need to tell the reader what Millican's argument is.

Here she is offering a supporting reason for accepting her claim that even logic (at least, in its current form) seems to take on board Kant's claim that existence is not a property. We like seeing this. Occasionally, students make bold claims without offering the reader any sort of reason for why this claim should be accepted. But as a general matter, you should be offering reasons to accept a given claim where you can. (Alternatively, if you are discussing a claim that you do not accept yourself, you should offer reasons that someone else has [or could have] offered for accepting their view.)

Again, per my earlier comment, I think the Author hasn't sufficiently explained to the reader what the Millican "irrelevance" objection is, and how it is supposed to work.

Anselm's argument that it is relevant to rather than showing why it is irrelevant to the others.

Here I shall object to the idea that the fool must accept (1) if he is to deny God's existence. By rejecting the existence of God, it is a mistake to think that the Fool automatically accepts P1. If he did, the argument would actually have to include an extra premise P0:

(P0) God does not exist. Anselm tells the Fool to believe that therefore:

(P1) God exists in the understanding but not in reality.

This missed step in AO uncovers how Kant's objection is relevant. According to Anselm, the Fool must accept P1. But if we lay out the argument more clearly in this way, we can see that the Fool could deny P1 (Heathwood 2011).

I will now show how Kant's objection is not irrelevant by looking at the problem of negative existentials. Anselm's move from P0 to P1 could be supported by Meinongianism (a solution to the negative existentials problem). Meinongianism argues that existence can be divided into two categories – existence in reality and also in the understanding. It is possible for things to exist just in the understanding, or to exist in both. (Meinong 1904:78-81) The tooth fairy, for example, exists only in the understanding. So, Meinongianism offers this as a solution to the problem of negative existentials. Anselm, by deducing P1 from P0, assumes a theory like Meinongianism (that there are two types of existence as a solution to negative existentials) and that in P0, God exists in the understanding but not in reality.

However, the problem is that if Meinongianism is true, then existence (in reality) must be a real predicate. I will now illustrate this using the example of the tooth fairy:

More signposting.

And more importantly, the author tells the reader exactly which premise of Anselm's argument she is rejecting. We love to see that.

More signposting.

This is worth highlighting.

Here, the author introduces crazy-sounding bit of philosophical jargon: 'Meinongianism'. This is not the sort of word that ordinary people can be expected to know. If the author does not explain what this word means, the reader is going to be lost. But the author **does** explain what the word means. So we are saved!...

...but no sooner does my heart break.

Here the reader refers to something called 'the problem of negative existentials'. I don't recall the author describing what "the problem of negative existentials" is. I imagine that most of you reading don't know what that is.

(Fortunately, the fact that Meinongianism offers a solution to the problem of negative existentials is not absolutely crucial to the author's argument. And given the space limitations, it might have taken her too far afield to try to explicate the problem of negative existentials. But I'm flagging this for you since, as a general matter, your philosophy professors won't like seeing you make reference to philosophical jargon without clarifying what it means. I should add that, I think a point like this could be stuck in a footnote. For instance, she could have put in a footnote: "Readers familiar with the problem of negative existentials might recognize Meinongianism as a solution to the problem.")

Again, the author here expands and deepens the explanation of her argument. This time she uses an illustration. Using concrete illustrations or examples are always a good idea when you are discussing abstract matters. This helps bring things down to a level where they are easier to grasp.

Consider concept A:

- is a fairy
- exchanges baby teeth for money
- exists in the understanding

And now consider concept A+:

- is a fairy
- exchanges baby teeth for money
- exists in the understanding
- exists in reality

In concept A+, the concept is changed from A, and so the thing that is added is a real predicate (from how we earlier defined a predicate).

Thus, Kant's objection to AO can be rewritten in premise and conclusion form (Heathwood 2011):

- P1* If AO is sound, then Meinongianism (or any theory similar) is true. (Anselm needs this to justify the step from P0 to P1 in his argument)
- P2* If Meinongianism (or any theory similar) is true, then existence in reality is a real property. (Tooth fairy example illustrates this)
- P3* But existence in reality is not a real property. (From Kant)
- C* Therefore, AO is not sound. (From P1, P2, P3)

In conclusion, I have argued that Kant's objection to Anselm's ontological argument (AO) does not succeed in proving the existence of God. I explained AO and then demonstrated Kant's objection that existence is not a real predicate using my own example of an orange ball. I offered a potential response to this, that Kant's objection is irrelevant to AO, but showed how this is not the case. This is due to there being a hidden premise in the standard way to set out AO into premise form and conclusion. It is this premise that Kant's objection is relevant to, which thus shows AO to be refuted.

A very brief conclusion that recaps the main points of the paper. It works for me.

References

Anselm, St. (1989). *The Ontological Argument*. W. I. Rowe and W. J. Wainwright (Eds.) *Philosophy of Religion: Selected Readings*. (2nd edition, pp. 96-98). New York: Harcourt Brace Jovanovich, Inc.

Heathwood, Chris. "The relevance of Kant's objection to Anselm's ontological argument." *Religious Studies* 47.03 (2011) pp. 345-357.

Kant, Immanuel. (1989). *Of the Impossibility of an Ontological Proof*. W. I. Rowe and W. J. Wainwright (Eds.) *Philosophy of Religion: Selected Readings*. (2nd edition, 105-110). New York: Harcourt Brace Jovanovich, Inc.

Meinong, Alexius. (1904) "The Theory of Objects" ("Uber Gegenstandstheorie"), pp. 78-81.

Millican, Peter. (2004) 'The one fatal flaw in Anselm's argument', *Mind*, 113, pp. 437-476.

External reference: it shows that the author went to the trouble of doing some research. (And she didn't just cite Heathwood to pad out her bibliography. She put the reference to good use.)

Another external reference. (If you are going to talk about Meinongianism, it makes sense to read a little bit from Meinong.)

Another external resource.

(We don't expect first year students to know this, but *Mind* happens to be one of the best philosophy journals in the English-speaking world [and arguably beyond]. But I want to take this opportunity to note that we like seeing you reference credible sources. Again, we know it's not easy for first year students to be able to tell what is a good journal from what isn't. But at the very least, citing work that has appeared in peer reviewed journals is much better than e.g., citing websites).

Rubin's Comment:

The blue comment bubbles are the comments I gave to the student on her essay as I was marking it. The yellow comment bubbles were added later to say a bit more about what she did right, and to flag some things that were less good (but which I didn't flag at the time).

PHIL1003: God, Mind and Knowledge

Roberta Burattini

Essay Question: Explain Descartes' version of substance dualism. Present and explain an objection to substance dualism. Can substance dualism overcome this challenge?

Introduction

~~The nature of the relationship between our minds and our bodies has been explored by many philosophers for centuries, and multiple~~ Many theories have been put forward in the attempt to understand ~~it: the nature of the mind-body relationship.~~ Among the most known, ~~is~~ Descartes' substance dualism ~~has gained many followers but also many critics, and it is debated upon even today.~~ In this essay I will first introduce ~~Descartes' substance~~ this dualism, then I will explain and evaluate ~~what's known as~~ 'the argument from causal closure' against it. I will conclude that the causal closure objection is not supported by modern quantum physics, and therefore, ~~e~~Cartesian dualism may overcome the challenge it presents.

Descartes' substance dualism

Descartes' substance dualism consists in the belief that mind and body are distinct, separated, things: while the mind is a mental (immaterial) substance, the body is a corporeal (material) substance (Descartes, 1642, p. 190). Starting from the fundamental principle "I think, therefore I am", Descartes establishes that he is a mind, a thinking thing, which can be conceived to exist without the body, a thing that he possesses (Descartes, 1642, p. 190). He then clarifies how the two substances interact in union: the mind can move the body, and the body can act on the mind through sensations (Princess E. of Bohemia et al., 2007, p. 65). As so far presented, Descartes' substance dualism assumes that a

Explain more about this.

Careful. This alludes to ideas and arguments that you don't have the room to explain. So better to avoid referring to them. It would be better to use these words to explain what Descartes means by mental and physical substance. What characterizes each? This is also what the essay question asked you to do.

non-physical substance can causally interact with a physical one. This assumption is challenged by materialism: the theory that the only existing substance in nature is matter. This theory objects substance dualism through the argument from causal closure. In what follows, I will briefly present this objection, then I will argue that modern quantum physics' discoveries do not support it.

Say a bit more about what this involves.

The argument from causal closure

The causal closure argument outlined as premises and conclusion (Papineau, 2002, pp. 17-18):

It would be good if you could have also briefly said in your own words what the gist of the argument is about. It would demonstrate that you understood it well.

P1: Conscious mental occurrences have physical effects

P2: All physical effects are fully caused by physical prior histories

P3: The physical effects of mental occurrences aren't always overdetermined by distinct causes

P4: Conscious mental events are physical events

C: Materialism is true

You haven't really explained this premise in the discussion below. Given that this is what you are challenging it's important to spell it out.

With this argument, materialism **attempts to prove** that Descartes' substance dualism is false since **it the latter** implies the existence of a non-physical substance that can cause physical events. I will now explain how recent discoveries in quantum mechanics **do not corroborate challenge** P2, and instead, point in the direction of a non-physical approach to causation of some physical events.

To say it proves it is to beg the question. ('Proves' is a success term).

Quantum mechanics and the observer effect

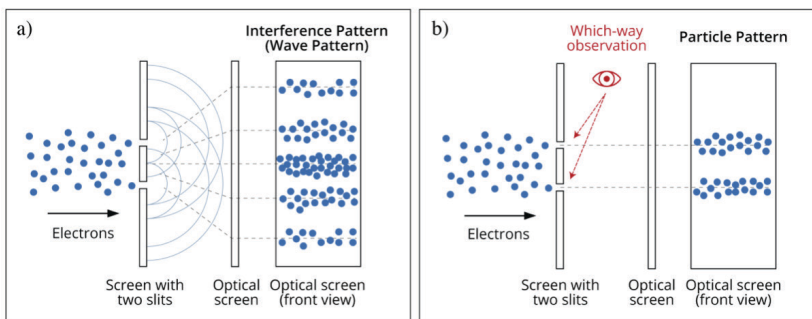
In quantum mechanics, the observer effect is the phenomenon by which a quantum wave collapses to present a specific outcome out of a range of possibilities. For the purposes of this essay, I will offer a simple explanation of this effect by introducing the double slit experiment that demonstrates it. In the double slit experiment, a series of electrons are sent in rapid succession through a surface with two slits. As shown in Figure 1 a), when there is no observation of which slit the electrons go through, they behave like a wave, showing an interference pattern (quantum wave function). If measurement is made, before or after electrons have passed the

slits, they act like particles: Figure 1 b). This experiment shows that an electron can act like a wave or like a particle depending on the fact that it is being observed (Narasimhan et al., 2019, p. 13). Before measurement, the electron is believed to be in a superposition: it is both wave and particle, it is in the realm of probability and has no definitive state (it has all the states at once), and only by observation it takes a determinate form.

Very nicely explained.

Figure 1

The Double Slit Experiment



The nature of the observer effect is still mostly unknown; many interpretations have been put forward, but no unique explanation has been formulated. Among the different theories, the orthodox interpretation of quantum mechanics (standard interpretation) seems the best equipped to find the answer to this mystery. I will use it to explain what implications it presents in regard to the argument from causal closure.

In the light of recent discoveries, it has been proposed that the explanation for the observer phenomenon can involve consciousness, defined as a non-physical force “rooted in a dimensionless space that is not subject to conventional limitations in space-time” (Narasimhan et al., 2019, p. 13). Researchers have come to this conclusion following the results of new experiments that show how the observer effect can happen outside of the space-time limits, in a manner that presupposes retro causality (Narasimhan et al., 2019, p. 15). They posit the existence of a “universal Observer, outside of what we perceive as classical

space-time, at an information level of existence” which has effects on the physical reality (Narasimhan et al., 2019, p. 16). This new evidence contradicts the second premise of the causal argument because: 1) it allows for events in the future to affect events in the past, 2) it identifies non-physical causes for physical effects.

How does reversing the temporal order undermine premise 2?

This more clearly challenges it.

~~In view of these outcomes, I argue that~~ If correct, this would put pressure on the causal closure argument, ~~undermining cannot be considered sound, therefore it loses its~~ efficacy as an objection to substance dualism.

Possible objection

What I have presented so far could be challenged by objecting that the observer effect is still a controversial topic with no widely accepted explanation. To this objection I can reply that even if there is not yet a unified understanding of the cause of the quantum wave’s collapse, this phenomenon cannot be interpreted or understood on classical physics terms. This is enough to doubt P2 and hence, ~~affect call into question~~ the soundness of the causal closure argument.

Conclusion

In this essay I have introduced only a small part of the abundant literature that connects quantum physics to the mind-body problem. Nonetheless, I believe that it is a good starting point for the revisitation of some objections that have been presented against substance dualism, allowing for its reconsideration as a compelling answer to the mind-body problem.

Good, nuanced conclusion.

Reference List

- Descartes, R., & Haldane, E. (1973). *Meditations on the First Philosophy*. In *The Philosophical Works of Descartes* (Vol. 1, pp. 140–199). Cambridge University Press.
- Narasimhan, A., Chopra, D., & Kafatos, M. C. (2019). The Nature of the Heisenberg-von Neumann Cut: Enhanced Orthodox Interpretation of Quantum Mechanics. *Activitas Nervosa Superior*, 61(1–2), 12–17. <https://doi.org/10.1007/s41470-019-00048-x>
- Papineau, D. (2002). The Case for Materialism. In *Thinking about Consciousness*. (pp. 13–21). Oxford University Press. <https://doi.org/10.1093/0199243824.003.0002>
- Princess Elisabeth of Bohemia, Descartes, R., & Shapiro, L. (2007). *The Correspondence*. In *The correspondence between Princess Elisabeth of Bohemia and René Descartes*. (pp. 61–73). University of Chicago Press. <https://doi.org/10.7208/9780226204444>

Miri's Comment:

This was a very good, well-written, well-structured and original essay that went out on a limb and for the most part did a nice job of challenging the causal closure argument in the few words that you had. The essay gave a very clear explanation of the double slit experiment along with its implications for the CC argument. I thought that the essay could have done a better job of explaining dualism near the beginning, as well as the causal closure premise that was the target of your critique. See other comments in the margin for feedback. On the whole, well done!

Level Two

PHIL2001: Bioethics

Tobias Langtry

Essay Question: Suppose a safe and effective COVID-19 vaccine had been developed. Would states be justified in compelling citizens to get vaccinated?

Introduction

This essay will consider whether it is morally justified for states to compel their citizens to be inoculated with a safe and effective vaccine that protects them against the infectious disease COVID-19. It will introduce the freedom from harm argument (FFHA) that attempts to justify compulsory vaccination. After explaining and providing support for the FFHA, this essay will introduce the declining harm objection (DHO). It will argue that the FFHA succeeds in justifying compulsory vaccination, but that this argument is weaker as the risk of transmission declines and is not successful when the risk is very low.

Work on grammar required.

Good intro — clear statement of intent.

The Freedom From Harm Argument

- P1) *Compulsory vaccination is morally justified if I) non-vaccination leads to a serious violation of another person's right to be free from unjustified harm or risk of harm, II) compulsory vaccination is sufficiently necessary and effective for preventing this violation, and III) the objective cost of being compelled to vaccinate is proportionate to the seriousness of this violation.*
- P2) *Non-vaccination leads to a serious violation of another person's right to be free from unjustified harm or risk of harm.*
- P3) *Compulsory vaccination is sufficiently necessary and effective for preventing this violation.*
- P4) *The objective costs born by being compelled to vaccinate are proportionate to the seriousness of this violation.*
- C) *Compulsory vaccination is morally justified.*
- (This is modified formulation of the argument made in Flanigan,

2014. The concept of 'objective cost' is borrowed from Giubilini, 2019).

Word choice error:
Sometimes choosing the correct word to express exactly what you have to say is very difficult to do. Word choice errors can be the result of not paying attention to the word or trying too hard to come up with a fancier word when a simple one is appropriate. A thesaurus can be a handy tool when you're trying to find a word that's similar to, but more accurate than, the one you're looking up. However, it can often introduce more problems if you use a word thinking it has exactly the same meaning.

The FFHA (formulated above) argues that compulsory vaccination is morally justified because it is legitimate for states to compel citizens to vaccinate in order to prevent those citizens from seriously violating the rights of other persons to be free from unjustified harm or risk of harm. This argument takes the form of the logically valid *modus ponens* deductive argument. Therefore, if all the premises are true then the conclusion must be true. This essay will assume premise three is true then will provide support for premises one, two and four. It will then explore the DHO which challenges premises two and four.

No need to explain *modus ponens* ... just make the claim that the argument is cogent via mp.

Work on grammar needed.
"... one, two and four, seeing if two and four can endure challenges raised by the DHO" ... or some such thing ...

Support for Premise One

The rationale empowering premise one is analogous to that which is commonly accepted as justifying laws which make random gunfire illegal. If there are no morally significant differences between compulsory vaccination and laws that make random gunfire illegal then it is morally justified for the state to make vaccination compulsory (Flanigan, 2014, pp. 7-10, 13-14).

Good.

Premise one is supported by our moral intuitions about when it is unjustified for the state to make vaccination compulsory or to make access to firearms illegal. State intervention would not be justified if an individual is likely to be harmed by vaccination (such as the ill or immunocompromised) or by lack of access to firearms. This is because the risk of harm to others is morally justified by the right to self-defence, or because the objective cost of state intervention is not proportionate to the seriousness of the violation (Flanigan, 2014, pp. 13-16).

I know you are working to a restrictive word limit, however both points needed further explanation.

Support for Premise Two

The non-vaccinated seriously violate another person's rights because the resultant harm is potentially very great and

unjustified. The non-vaccinated do actual or probable harm directly by increasing the risk that they will be infected and then spread COVID-19 to other persons. They also do indirect harm by undermining herd immunity, increasing the risk of very large infection chains and by limiting access to valuable social goods for those who are vulnerable to infection (Flanigan, 2014, pp. 6, 9; Umbers, 2020A, pp. 4).

This harm is morally unjustified because most probable victims do not revoke their relevant rights (by providing consent), and most of the non-vaccinated are not justified in refusing to vaccinate because the objective cost to them is insufficient (Flanigan, 2014, pp. 12, 14-15).

Support for Premise Four

Even safe vaccines pose the very low risk of severe health complications (illness, death). This does not however imply that the costs of compulsory vaccination are in most cases disproportionate to the seriousness of the violations prevented. Just as laws that require the use of seatbelts are proportionate, despite the low risk of severe spinal injury, compulsory vaccination is proportionate because the risk of a severe complication is sufficiently low (Giubilini, 2019, pp. 456-458).

Good examples.

Making vaccination compulsory poses the risk of this policy backfiring by undermining trust in the medical community and vaccines. However, this cost is not sufficient to render compulsory vaccination disproportionate because evidence suggests that compulsory vaccination when implemented properly is likely to increase rates of vaccination with minimal backfiring (Giubilini, 2019, pp. 462-463).

The Declining Harms Objection

The DHO challenges premises two and four of the FFHA. It argues that the strength of these premises declines as the risk of

transmission declines and that these premises are not true when the risk is very low, such as when herd immunity has been reached.

Herd immunity refers to the significant reduction in rate and risk of transmission that occurs when a sufficient proportion of a population is immunised against a transmissible virus. As the risk of transmission declines the risk of the non-vaccinated being infected and spreading COVID-19 declines. The actual or probable harm resulting from non-vaccination declines. The seriousness of the non-vaccinated person's violation of another person's right declines and the cost of compulsory vaccination is less proportionate. If this risk is very low premises two and four are no longer true (Giubilini, 2019, pp. 448-449).

This statement was needed at first mention of herd immunity.

A Response to the Declining Harms Objection

Jessica Flanigan responds to this objection by pointing out that even when the non-vaccinated are not likely to do direct harm because the risk of transmission is low, they do contribute to an aggregate harm. They do so by undermining herd immunity, by lightly increasing the risk of transmission and by potentially encouraging others to refuse vaccination. When a significant proportion refuses to vaccinate this aggregate harm can be very great as the risk of COVID-19 transmission rises significantly. However, the contribution each person makes is usually minimal and is only significant when many people refuse to vaccinate (Flanigan, 2014, pp. 11-12; Giubilini, 2019, pp. 448-449).

Always refer to authors by their surname only ... Flanigan (2014) ...

Contrary to the DHO, premise two is true because the contribution the non-vaccinated make to this aggregate harm is a serious violation of another person's right to be free from unjustified harm or risk of harm. Premise four is true because the cost of vaccination is proportionate to the seriousness of this violation. Therefore compulsory vaccination is justified even when risk of transmission is low.

Good.

Does this Response work?

This response is insufficient. The FFHA assumes that the morally significant feature of non-vaccination that justifies compulsory vaccination is how seriously one violates another person's rights. However, by contributing to an aggregate harm an individual person does not seriously violate any person's rights. This is because this aggregate harm is over-determined and would occur even if one additional person was vaccinated. It is extremely difficult to identify who was harmed by whom and how seriously their rights were violated, if at all (Umbers, 2020B, pp. 4-5). The FFHA cannot therefore justify compulsory vaccination by reference to preventable aggregate harms.

Even when there is a non-arbitrarily identifiable victim caused by non-vaccination when the risk of transmission is very low, the seriousness of one's violation is not likely to be serious enough to justify compulsory vaccination, precisely because the likely risk of being infected and spreading COVID-19 is very low (Giubilini, 2019, pp. 449).

Word choice — consider changing.

Therefore, when the risk of transmitting COVID-19 is very low the FFHA fails to justify compulsory vaccination because premises two and or four are factually incorrect. If compulsory vaccination is still justified when the risk of transmission is very low, it is not by reference to rights to be free from harm of risk of harm.

Conclusion

The essay started by explaining and supporting the FFHA, which justifies compulsory vaccination on the grounds that compulsory vaccination prevents individuals from seriously violating the rights of another person to be free from unjustified harm or risk of harm. It then sketched out the DHO response to the FFHA which argued that the FFHA fails when the risk of transmitting COVID-19 is low. This is because the risk of doing harm by being infected and spreading COVID-19 is much lower. It then argued that while the

non-vaccinated contribute to a potentially great aggregate harm they do not violate another person's rights seriously enough. The FFHA therefore is able to justify compulsory vaccination except where the risk of transmission is very low and the strength of the FFHA declines as the risk of transmission does.

References — APA

- Flanigan, J. (2014). A Defense of Compulsory Vaccination. *HEC Forum*, 26(1), 5-25.
- Giubilini, A. (2019). An Argument for Compulsory Vaccination: The Taxation Analogy. *Journal of Applied Philosophy*, 37(3), 446-466.
- Umbers, L. (2020A). *Bioethics: Ethics & COVID-19 II – Lockdown* [Powerpoint slides]. PHIL2001, University of Western Australia, Australia.
- Umbers, L. (2020B). *Bioethics: Ethics & COVID-19 III – Vaccination* [Powerpoint slides]. PHIL2001, University of Western Australia, Australia.

Kaz's Comment:

An excellent, well argued response. A structural suggestion: some initial outlining of key terms, such as compulsory vs. mandatory vaccination and herd immunity in particular, would have set you up for the essay. Your arguments are concise and to the point, however there are places where more explanation was needed. Otherwise a really really good submission. Congratulations!

PHIL2006: Philosophy of Psychology and Psychiatry

Kevin Fitzgerald

Essay Question: Classicism. Describe and critically evaluate the classical computational theory of mind.

This paper will seek to describe the fundamental tenets of the classical computational theory of mind (henceforth referred to as 'CTM'), and critically evaluate its viability. Owing to its status as the dominant theory of mind during the 1980s, CTM can more accurately be described as a "family" of views, rather than a single specific formulation. However, the most prolific advocates of CTM such as Jerry Fodor¹ and Zenon Pylyshyn² provide their own formulations of the view. For the purposes of this paper I will explore Fodor's version of CTM, which he calls *the representational theory of mind (RTM)*. Additionally, I will consider the cogency of several objections to the CTM, such as John Searle's infamous "Chinese Room" thought experiment³ (1980), and alternative paradigms such as connectionism. I will conclude that CTM no longer provides the strongest theory of mind.

To understand the fundamentals of CTM and the various critiques that followed in any substantive way, it will be necessary to dedicate some time in this paper to provide a brief overview of the key concepts and terms within the CTM school of thought. Philosophers of mind make an important distinction between *semantics* versus *syntax* within a symbol system (also known as a 'formal system'). The syntax of a symbol system refers to the *rules* by which the symbols may be manipulated, combined or moved within the system. The semantics refers to the 'meaning-laden' or 'meaning-involved' properties of a given symbol. In other words, what the symbol means within the world. Within a given formal system (for example, chess or tic-tac-toe), if the syntax is designed and observed properly, symbols should always be able to be manipulated in such a way that respects its semantic content. In an *interpreted automatic* formal system, the system runs by itself and interprets symbols by following the syntactic rules it has been provided. Alan Turing's 'Turing Machine'⁴ demonstrated that such a system was indeed possible to create, the unique insight being that such a syntax governed system would not be limited to narrow mathematical operations (such as a calculator), but could solve any well-specified

¹ (1975, 1980a, 1981, 1987, 1990).

² (1980, 1984).

³ Searle, J – *Minds, Brains and Programs* (1980), *The Behavioural and Brain Sciences*.

⁴ See, Turing, A.M – *On Computable Numbers, With an Application to the Entscheidungsproblem*. (1936-7).

problem. This inspired many (including Turing) to work towards creating “thinking machinery” that could execute core mental tasks such as reasoning, decision-making and problem solving. A system of mechanical ‘formal symbol manipulation’⁵ as we have described, is the basis for how a digital computer operates. Proponents of CTM argue that this is a viable framework for how we can, in a literal sense – understand the mind: as a computational system. Mental states are held to be “representational”. This is to say, that mental states are comprised of symbolic representations that have both semantic and syntactic properties.

The foundations of CTM are thus built on two main thesis:

- 1) CTM holds that intentional states are relational states involving a thinker and symbolic representations of the content of the states. (*an account of mental states*)
- 2) CTM holds that cognitive processes are computations over mental representations. (*an account of reasoning*)

Jerry Fodor describes the system of mental representations as occurring in a ‘language of thought’ (referred to as *mentalese*)⁶. He posits the existence of a system of mental representations comprised of primitive representations and complex representations (formed through combining primitive representations together). He gives the example of how the primitive mentalese words JOHN, MARY, and LOVES can be combined to form the mentalese sentence JOHN LOVES MARY – demonstrating that the meaning of complex mentalese expressions are a function of the meaning of its constitutive parts and how they are combined. Additionally, he presents intentional states (also referred to as a propositional attitude, I will use these interchangeably) as *relations* to these Mentalese symbols. Intentional states can be understood as a type of mental state that has the characteristic of *being about something or being directed towards something*. *Fearing* that there is a dog behind the fence, as opposed to *hoping* there is a dog behind the fence is an example of two different functional relations (sometimes characterised as an *attitude*) to a symbolic mental representation (the dog behind the fence) with the same semantic value. This is to be distinguished from other descriptions of mental phenomena such as ‘qualia’⁷, that are not directed towards or about anything.

Cognitive activity, according to Fodor, can be explained as a formalised computational process over a language of thought. Mental computation occurring in the brain would store mentalese symbols in physical memory locations and manipulate these symbols according to syntactic rules. Such a framework has two significant appeals. Firstly, the schema of a language of thought allows for a finite

⁵ Henceforth referred to as “formalisation”.

⁶ Fodor, J. A – *The Language Of Thought* (1975).

⁷ See, Chalmers, D – *The Conscious Mind* (1996).

set of primitive Mentalese expressions to generate an infinite set of complex mentalese expressions. This allows for, in theory, the instantiation of an infinite set of intentional states. We will refer to this generative ability as *productivity*. Secondly, it seems to provide an account for the systematic relations between mental states, or *systematicity*. Someone who understands the phrase “the fox chased the hare” would also be able to understand the sentence “the hare chased the fox” (as unlikely as that scenario may be in the real world). CTM’s ability to account for productivity and systematicity demonstrated congruence with important empirical research of its time, such as Noam Chomsky’s “cognitivist revolution” within the field of linguistics⁸. Chomsky argued that the behaviourist paradigm⁹ was unable to explain how children observed and latched onto grammatical rules, and were then able to apply those rules in an indefinite amount of novel contexts. Children would have to apply these rules in ways undetermined by the finite set of language they have been exposed to. This language learning process by a child under the Chomskian model was characterised as a process of formation and confirmation of hypotheses, and as such, according to Fodor – would require such a language of thought to exist.

Moreover, CTM seems to also provide an explanation for how reasoning functions as a casual process. Formalisation gives us the ability to link semantics to syntax, and computation seems to show a physical mechanism by which we can link syntax to a casual process that respects the semantic values of the terms. This construction of the mind as a “syntactic engine”¹⁰ convinced many philosophers of mind that CTM was the “only game in town”. However, after enjoying its status as the dominant theory throughout the 1980s and 1990s, alternative paradigms for explaining psychological process such as “connectionism”¹¹ have emerged.

No critique of CTM has been as widely discussed as John Searle’s infamous “Chinese Room” thought experiment¹². For the sake of brevity, I will provide the most concise summarisation directly from Searle two decades after he originally published his argument:

Imagine a native English speaker who knows no Chinese locked in a room full of boxes of Chinese symbols (a data base) together with a book of instructions for manipulating the symbols (the program). Imagine that people outside the room send in other Chinese symbols which, unknown to the person in the room, are questions in Chinese (the input). And imagine that by following the instructions in the program the man in the room is able to pass out Chinese symbols which are correct answers to the questions (the output). The program enables the person in the room to pass the Turing Test

⁸ Chomsky, N – *A Review Of B.F Skinner* (1959).

⁹ Behaviourism sought to explain human and animal behaviour in terms of external physical stimuli, responses, learning histories, and reinforcements.

¹⁰ Haugeland, J – *Semantic engines: An Introduction to mind design*. (1981).

¹¹ See, Rumelhart & McClelland – *Parallel Distributed Processing* (1986).

¹² Searle, J – *Minds, Brains and Programs* (1980), *The Behavioural and Brain Sciences*.

for understanding Chinese but he does not understand a word of Chinese.¹³

Searle's experiment is constructed as a direct answer to Alan Turing's "Turing Test". Turing stipulated that we could successfully determine if a computer was truly a 'thinking machine' if a panel of questioners, on the basis of the answers it received alone – could not distinguish between the answers given by a person or a machine. The force of Searle's argument relies heavily on intuition, with Searle clearly hoping that readers side with his conclusion: that we would not characterise the native English speaker inside the room as someone who 'understands' Chinese, even if that person is successfully fooling the Chinese speakers outside the room. Ostensibly, Searle's argument is that mere simulation of linguistic competence – through the strictly computational means CTM endorses, does not provide the sufficient conditions for *understanding* of meaning or semantics. Ned Block¹⁴ characterises the Chinese Room's forceful reliance on intuition as a weakness, arguing that potentially there is a need to revise our intuitive concept of understanding, however – there is insufficient room in this paper to address all the replies to Searle, so I will address what I believe to be the most interesting one: the Robot Reply.¹⁵

The Robot Reply concedes that perhaps Searle is right: the man in the room doesn't understand Chinese. However, if we were able to incorporate our symbol manipulation system (digital computer) into a robot with sensors, microphones, wheels and other perceptual inputs and effects that allow it to interact and move around in the world, this would allow it to attach meaning to symbols and understand natural language in the semantic sense. This requirement for a sort of 'grounding' in the external world has appeal and demonstrates congruence with work being done in developmental robotics. However, unfortunately for advocates for CTM, it concedes something fatal for the entire project: "thinking cannot be simply symbol manipulation".¹⁶

Beginning in the late 1980s a rival paradigm emerged, challenging Fodor's belief that CTM was the "only game in town". Connectionism as a movement within cognitive science hoped to challenge CTM and explain intelligence through artificial neural networks. Rather than information being stored symbolically, the connectionist claims that it is stored non-symbolically in the weights, or connection strengths, between the units of a neural net. While the highly technical mechanics of connectionism cannot be explored *in-depth* in this paper, we can form a sketch of one of its primary strengths. Connectionism provides a more viable account for explaining Moravec's paradox; the observation in

¹³ Searle, J – 'The Chinese Room', in R.A. Wilson and F. Keil (eds.) (1999), *The MIT Encyclopedia of the Cognitive Sciences*, Cambridge, MA: MIT Press

¹⁴ Block, Ned – *Readings in Philosophy of Psychology*, 2 vols. Vol. 1. (Cambridge: Harvard, 1980).

¹⁵ Versions of this reply have been endorsed by Margaret Boden, Tim Crane, Daniel Dennett – amongst others.

¹⁶ Crane, T – *The Mechanical Mind* (1991).

AI research that “the hard problems are easy and the easy problems are hard”.¹⁷ While classical AI has little difficulty in exhibiting traditional notions of intelligence to solve the ‘hard problems’ such as playing chess or solving complex mathematical equations (things that are understood as characteristic of a highly intelligent humans), it struggled to learn and solve the ‘easy’ problems such as perceptual processing (recognising faces) or motor control (walking around a room, shooting a basketball). Additionally, Connectionist systems exhibit *graceful degradation*. While a single error or failure within a classicist model can lead to failure of the entire system, connectionist models can be ‘lesioned’ or have their performance impaired proportionate to the extent of the lesion. Connectionism seeks to give an account of the “flexibility and insight found in human intelligence using methods that cannot be easily expressed in the form of exception free principles”¹⁸, allowing it to avoid the fragility that arises from standard forms of symbolic representation.

It is clear that while many dispute the formulation and the corresponding intuitions derived from Searle’s Chinese Room, it has remained one of the most discussed and thought provoking contributions to the philosophy of mind. For advocates of CTM, much turns on whether a computational process in and of itself, is sufficient for understanding. The challenges posted by Searle’s Chinese Room, combined with the increasing prevalence and complexity of neural networks has effectively put the classicist paradigm on the defensive, where it has arguably remained ever since. It seems clear that classical computational theorists of mind have much work to do to reclaim the legitimacy of the strictly rule-based symbol manipulation and computation as the dominant paradigm.

¹⁷ Pinker, S – *The Language Instinct* (1994), *Perennial Modern Classics*, Harper.

¹⁸ Horgan & Tienson – “Representations without Rules”, *Philosophical Topics*, 17(1): 147–174. (1989,1990).

REFERENCES:

Block, Ned – Readings in Philosophy of Psychology, 2 vols. Vol. 1. (Cambridge: Harvard, 1980).

Chalmers, David., 1996, *The Conscious Mind*, Oxford: Oxford University Press.

Chomsky, Noam – A Review Of B.F Skinner (1959).

Crane, Tim – *The Mechanical Mind* (1991).

David E. Rumelhart & James L. McClelland – *Parallel Distributed Processing* (1986). ISBN:

9780262181204

Pinker, Steven – *The Language Instinct* (1994), Perennial Modern Classics, Harper. ISBN 0-203-

42631-2

Searle, John – ‘The Chinese Room’, in R.A. Wilson and F. Keil (eds.) (1999), *The MIT Encyclopedia of the Cognitive Sciences*, Cambridge, MA: MIT Press.

Turing, Alan .M., 1936–7, “On Computable Numbers, With an Application to the

Entscheidungsproblem”, *Proceedings of the London Mathematical Society*, s2-42: 230–265

Level Three

PHIL3003: Moral Theory

Anonymous

Essay Question: Defend a moral theory from an objection.

1. Introduction

What is the right thing to do? What makes an act right? In this essay I explicate act consequentialism, one ethical theory attempting to answer these questions. By refuting common intuitions about causation, I also respond to an objection against act consequentialism presented by Richard Brandt in *Moral Obligation and General Welfare*.

The point you are making here should be discussed in greater depth.

(We recognise that, in some cases, you might find that it isn't possible to expand, due to the word limit. In any case, if you could find a way to free up some space, this is a part of the essay that could do with more attention.)

It's good to keep it short. But it would be good to give the reader just a bit more of a preview than this. Maybe two more sentences: e.g., roughly what's the problem with causation? What's Brandt's problem?

2. Act-Consequentialism

Act-consequentialism claims that an act is right if and only if it maximises utility. This means that out of all possible alternatives, it produces the greatest amount of good minus bad. The intrinsic nature of the act is irrelevant. Whether an act includes lying or stealing only matters to the extent it affects the consequences.

Insert a comma here.

Something here is imprecise. "at least as great"

Different act-consequentialist theories can adopt differing definitions of what is good/valuable. Hedonism holds pleasure as the only intrinsic good and pain as the only intrinsic bad (Mill, 1864, p. 234). Hedonic act-utilitarianism (AUh) therefore claims that the act maximising hedonic utility, the total amount of pleasure minus pain, is right. For simplicity, hedonic utilitarianism will be the act-consequentialist theory adopted in this essay.

3. An Objection to Hedonic Act-Utilitarianism

Richard Brandt (1959) presents the following thought experiment.

Case 1: Joe is an Englishman living during World War II. To preserve resources necessary to win the war, the English government requests that each citizen use a maximum ten litres of gas per week. However, Joe uses twenty litres. He

justifies this by claiming that it has a negligible effect on the war yet greatly improves his personal comfort.

While Brandt doesn't make a distinction, there are two contexts in which Joe's act can be evaluated. In context (a), the war is unaffected because no other Englishmen use the extra gas. In context (b), the war is lost because many Englishmen adopt Joe's thinking and use the extra gas. We can now consider the first of two arguments Brandt constructs against act utilitarianism.

- P1: If hedonic act-utilitarianism is correct, then Joe's act is right in context (a).
 P2: Joe's act is not right in context (a).
 C: Hedonic act-utilitarianism is not correct.

As Joe benefits from his act, and it doesn't harm anyone else, it maximises hedonic utility. AUh must therefore say it's right (P1). Premise two is motivated purely by our moral intuitions.

However, some act-utilitarians may deny these intuitions. Brandt therefore presents a stronger second argument.

- P1: If hedonic act-utilitarianism is correct, then every Englishman who uses extra gas in context (b) acts rightly.
 P2: Every Englishman who uses extra gas in context (b) does not act rightly.
 C: Hedonic act-utilitarianism is not correct.

Two arguments could support premise one. Firstly, if act-utilitarianism says Joe's act is right in context (a), Brandt argues the same act must also be right in context (b). Secondly, it may be argued that each Englishman causes no harm. After all, if he didn't use the gas the war would still be lost. As the extra gas benefits him, AUh must therefore claim the act is right. Walter-Sinnott Armstrong (2005) makes a similar argument.

However, while individual acts don't appear to cause harm,

Needs some additional explanation.

Although it may be implicit in what you have written in this sentence (or paragraph), there is a detail that should be made explicit for your reader.

Right. And so, the total aggregate utility in scenario (b) is higher if he used the gas, than it would be if he didn't.

Your citation should contain a reference the page where you found this quotation or idea. Not all citations must specify a specific page. But if there is a particular idea from a text that you are making reference to, then it's usually best to give the page number. (When quoting a passage, it's essential to give the page number.)

together the Englishmen's acts cause the war to be lost. Every act leading to such public harm must be wrong (P2).

4. Responding to Brandt's First Argument

I deny premise two in Brandt's first argument. Joe's act in context (a) is right. As this essay is mainly concerned with Brandt's second argument, I will not vigorously defend this claim. Nevertheless, two points must be considered. Firstly, our moral intuitions may be influenced by our disgust for Joe's selfish character. However, judgements of his character must be separated from judgements of his act. Secondly, Joe could be a 90-year-old war veteran with a broken hip, fractured leg and living alone in an uninsulated apartment. If the use of gas doesn't cause harm, surely we couldn't refuse Joe's desire for warmth at night. While not all Englishmen will be in such positions, this demonstrates that our moral intuitions are not as clear as Brandt suggests.

I have a misgiving about this point or this word/phrase.

I agree that these things seem morally relevant. But Brandt surely had in mind cases where the Englishman doesn't have any special maladies. And the argument would remain interesting if we simply stipulated that we were thinking of such a person.

5. Responding to Brandt's Second Argument

In support for premise one of Brandt's second argument it was claimed that individual acts of gas overuse cause no harm. By rebutting this, I deny premise one: that act-utilitarianism must say every Englishman using extra gas in context (b) acts rightly.

Imagine one million Englishmen decide to use extra gas at the same time and that this results in the war being lost. Who is responsible for causing this harm? "Not me!" One Englishman may claim, "for the war would have been lost regardless of my action." This man's argument appears plausible. However, if this man was right, then every Englishman could adopt this reasoning. This would lead to the paradoxical situation where no one and nothing is responsible for causing the harm. The Englishman has made Derek Parfit's second mistake in moral mathematics – to assume an act doesn't cause harm because the harm would have occurred anyway (Parfit, 1986, p.4). The paradox can only be avoided by holding every Englishman partly responsible for causing the harm. Hence, the second supporting argument for premise one fails.

I have a misgiving about this point or this word/phrase.

Even so, AU still has a problem: his act of using the gas is still the one that maximizes utility. It's not yet clear that it's being a partial or contributing cause of a great harm changes the fact that no other alternative for the individual would have resulted in a higher utility.

The first argument supporting premise one – that if Joe’s act is right in context (a), then it must be right in context (b) – also fails. Because Brandt doesn’t recognise the changing contexts, he also fails to recognise how the consequences of the same act change in these contexts. While no harm is produced in context (a), in context (b) every Englishman is responsible for causing harm.

Furthermore, this harm outweighs any benefits. If one million Englishmen use the extra gas, each man causes 1/1,000,000 of the harm from the war’s loss. If there were 40 million people living in England, each man would then be responsible for causing the harm for 40 of these people. This would clearly outweigh the benefits from the extra gas. Act utilitarianism will therefore say every Englishman using extra gas act wrongly. Premise one of Brandt’s second argument is false.

6. Potential Objections and My Responses

The opponent of act-utilitarianism may point out that the Englishmen would not use the gas simultaneously. Furthermore, the war may be lost only after a ‘threshold’ number of Englishmen use the extra gas. As a result, only the person who crosses this threshold causes harm (Fragniere, 2016, p. 801).

Even if the harms are produced in this threshold fashion, we must consider why the Englishman who crosses the threshold causes harm. This man causes harm only if many Englishmen have previously used extra gas. These previous acts have positioned the Englishman on a threshold point. Every Englishman who previously used extra gas therefore causes this man to cause harm. These Englishmen are therefore still responsible for causing harm, even if indirectly. Act-utilitarianism hence overcomes this objection.

A second objector could claim the harm is emergent (Kingston & Sinnott-Armstrong, 2018, p. 175). Consider a puddle of oil. While this puddle is slippery, no individual oil particle is itself slippery. Slipperiness is an emergent property: it belongs to the group but no individual particle. Similarly, the harm from the war’s loss could also be emergent. Perhaps the collection of actions (of

I have a misgiving about this point or this word/phrase.

I think this is a point well worth making! But I'm a little worried about signing on. After all, the agent at the threshold is not compelled by me to do their act. They might see that their act would throw things over the threshold, and on that account, refrain from doing it. But if that could happen, then it's not clear I caused them to do anything.

extra gas use) can be harmful while no individual act is itself.

However, if the harm was emergent, each individual act would still indirectly cause harm. Why? Because every individual act is partly responsible for creating the collection of acts. If we denied this then we would once again produce the paradox where no act (and nothing) is responsible for creating the collection of actions. As each individual act of extra gas use must be creating the collection of actions, every individual remains partly responsible for causing the harm that the collection of acts produces.

A third objector may claim my account adopts what Derek Parfit calls the 'share-of-the-total-view'. If a group acts to produce some particular outcome, the share-of-the-total view holds each individual responsible for producing the fraction of the outcome in accordance with the fraction of the group they represent (Parfit, 1986, p. 2). Similarly, in section six I suggested that if one million Englishmen use extra gas, each causes 1/1,000,000 of the harm. The following case demonstrates how this view produces incorrect results.

Case 2: Four people are required to save 100 miners trapped in a mining shaft. One person is required to save 10 miners trapped in a different shaft. Four rescuers are already working to save the 100 miners. Should you assist these four men or save the other 10 miners?

While we intuitively think we should save the ten miners, my previous argument appears to suggest that you should assist the four rescuers. This is because, by helping the rescuers, your act is responsible for causing 20 of the 100 lives to be saved. Furthermore, saving twenty lives is better than saving ten.

However, this case is disanalogous to Brandt's initial case. In Parfit's case you are making a decision after the four rescuers have started saving the 100 miners. In Brandt's case, Joe makes his decision before the war has been lost. We must therefore

distinguish between decisions made before and after a particular outcome is guaranteed.

My claim that each individual partially causes the harms only applies when their decision or action is made before the outcomes are guaranteed. My view therefore advises you to save the ten miners. Because the 100 miners are already guaranteed to be saved, by helping the rescuers you wouldn't be responsible for saving any lives. Similarly, if the war was already lost, my view suggests that Joe should use the extra gas. After-all, if the war was already lost, Joe's act wouldn't cause any harm.

My view avoids the problems outlined by Parfit's case by not applying the share-of-the-total-view in these circumstances. However, I simultaneously maintain that Joe causes harm in Brandt's case because the loss of the war isn't guaranteed when he uses the extra gas. Only such contexts where an outcome isn't yet guaranteed does my view adopt the share-of-the-total-view.

Finally, the opponent of act-consequentialism may claim my argument proves too much. On my view even those who use less than the prescribed amount of gas cause harm. This is the case. Once again, denying this would produce the paradox described above. However, these people used less gas. Therefore, they are responsible for less of the harm. Furthermore, they are likely using it only for essential services such as cooking. The benefits from this, like not starving, will likely outweigh the minimised harm. In cases where the gas is used for non-essential services such as leisurely drives, the harm may outweigh the benefits. Such acts would, and should, be classified as wrong.

I wonder if some appeal to expected utility (relying on subjective probability) isn't laying behind this thought.

Orthodox consequentialism won't really be sensitive to what the agent might reasonably anticipate will be the consequences of a given action.

It wouldn't be bad to appeal to expected utility. But it's better if it's explicit.

7. Why Brandt's Argument Failed

Premise one in Brandt's second argument fails because it adopted the counterfactual theory of causation. This theory proposes that "some event is causally efficacious only if, had it not occurred, an outcome would not have occurred just as it did" (Shafer-Landau, 1994, p. 86). This suggests that Joe did not cause harm in context (b) because the war would be lost regardless of his actions.

However, I have refuted this argument by demonstrating how the counterfactual theory of causation fails in cases where harms are overdetermined. This is because it leads to the paradox where there is harm, yet no one or nothing is responsible for creating that harm. The only solution to this paradox is to abandon the counterfactual theory and hold every individual partly responsible for causing the harm.

I have a misgiving about this point or this word/phrase.

Maybe that was behind your responses. But I don't remember an explicit discussion of the CF theory (you never really paused to lay it out), or of its rivals.

8. Conclusion

Given the failure of the counterfactual theory, a new theory of causation must be adopted to explain overdetermined harms. However, developing this theory is beyond the scope of this essay. Instead, this essay established how act-utilitarians can reject premise two in Brandt's first argument and premise one in his second argument. In context (a) Joe's act is right. However, in context (b), the same act is wrong. If any Englishman uses extra gas and the war is later lost, he becomes responsible for causing the harm, even if the war would have been lost anyway. Act-consequentialism can hence overcome and provide a satisfactory answer to the objection presented by Richard Brandt.

References

- Brandt, R. (1959). *Ethical Theory: The Problems of Normative and Critical Ethics*. New Jersey: Prentice-Hall, Inc.
- Fragnière, A. (2016). Climate change and individual duties. *Wiley Interdisciplinary Reviews: Climate Change*, 7(6), 798–814. <https://doi.org/10.1002/wcc.422>
- Kingston, E., & Sinnott-Armstrong, W. (2018). What's Wrong with Joyguzzling? *Ethical Theory and Moral Practice*, 21(1), 169–186. <https://doi.org/10.1007/s10677-017-9859-1>
- Mill, JS. (1969). *Utilitarianism*. Toronto: University of Toronto Press.
- Parfit, D. (1986). *Five Mistakes in Moral Mathematics*. <https://doi.org/10.1093/019824908X.003.0003>
- Schwenkenbecher, A. (2014). Is there an obligation to reduce one's individual carbon footprint? *Critical Review of International Social and Political Philosophy*, 17(2), 168–188. <https://doi.org/10.1080/13698230.2012.692984>
- Shafer-Landau, R. (1994). Vegetarianism, Causation and Ethical Theory. *Public Affairs Quarterly*, 8(1), 85–100. Retrieved from <https://www.jstor.org/stable/40435869>
- Sinnott-Armstrong, W. (2005). It's not My Fault: Global Warming and Individual Moral Obligations. In W. Sinnott-Armstrong & R. B. Howarth (Eds.), *Perspectives on Climate Change: Science, Economics, Politics, Ethics* (1st ed., pp. 285–307). Burlington: Emerald Group Publishing.
- Rosenqvist, S. (2019). The No Act Objection: Act-Consequentialism and Coordination Games. *Thought: A Journal of Philosophy*, 8(3), 179–189. <https://doi.org/10.1002/tht3.418>

Rubin's Comment:

There are things that could be tightened up in the presentation. But overall, I think this is a great discussion. I really appreciated the last few pages with objections and replies. It helped clarify things for me a lot.

PHIL3003: Moral Theory

Aidan Mansfield

Essay Question: Describe a theory of well-being that was either discussed in class, or else has been held by a philosopher in an academic publication. Offer a critique of that theory.

I will argue that hedonism fails as a theory of wellbeing due to the immeasurability of pleasure and pain. To demonstrate this failure, I first outline a basic theory of hedonism as formulated by Mill. Secondly, I explain how the immeasurability of pleasure and pain creates problems for this hedonism. Thirdly, I discuss an objection to this critique. Finally, I attempt to offer a response towards this objection and conclude that hedonism fails as a successful theory of wellbeing.

Of wellbeing, Mill argues “happiness is desirable, and the only thing desirable, as an end; all other things being only desirable as a means to that end”¹ Mill equates happiness with pleasure,² and in suggesting that these are the only things desirable as an end, argues pleasure is the sole basic bearer of intrinsic value, and pain is the sole basic bearer of intrinsic badness.³ Based on this, we can construct a basic argument for hedonism (H).

1. Every instance of pleasure is intrinsically good;
Every instance of pain is intrinsically bad;
Nothing else has basic intrinsic value
2. The intrinsic value of an instance of pleasure = (the intensity of the pleasure) (its duration);
The intrinsic value of an instance of pain = (the intensity of the pain) (its duration)
3. The intrinsic value of a composite thing is equal to the sum of the intrinsic values of all the instances of pleasure and pain that it contains.⁴

Accordingly, hedonists regard a life with the greatest proportion of pleasure over pain the most intrinsically valuable, prescribing actions which lead to maximising the ratio of pleasure over pain. To identify these actions, a subject must be able to rank experiences of pain and pleasure in a coherent

¹ John Stuart Mill, (1861), Utilitarianism, reprinted in *The Collected Works of John Stuart Mill, Volume X – Essays on Ethics, Religion and Society*, ed. John M. Robson, Introduction by F. E. L. Priestly (Toronto: University of Toronto Press, London: Routledge and Kegan Paul, 1969), 234.

² Ibid, 210.

³ Michael Rubin, “Subjective Theories of Value: Hedonism” (Lecture, University of Western Australia, Crawley, WA, March 9, 2020).

⁴ Ibid.

hierarchy. Mill claims this can be done through someone who has experienced both pleasures, giving one a “decided preference”.⁵ When a critical eye is cast over the second premise, however, it seems unintuitive to suggest that pleasure comes in amounts exact enough to produce a ‘decided preference’. It seems intuitive to say we draw different forms of pleasure from different things; pleasure from intellectual pursuits, love, friendship, art and sport are all fundamentally different in nature.⁶ Hedonism, it seems, demands a kind of “homogenising” of pleasure,⁷ that seems highly reductive.

The intuition that pleasure comes in fundamentally different forms can be explored by counterexample. Consider a subject who gains substantial pleasure from both the activity of travelling, and of spending time with friends. It seems reasonable to say, if questioned, the subject would struggle to produce an answer as to which activity produces the greater pleasure. The subject would likely conclude the pleasures are so similar, yet fundamentally different, thus incomparable. So surely, we cannot precisely calculate the intrinsic value of instances of pain and pleasure such to produce a ‘decided preference’.

Franz Brentano’s counterexample is more developed. He explains there is no “number, n , such that it would be right to say that one episode contains exactly n times as much pleasure as another”.⁸ To demonstrate this, Brentano asks the hedonist to “consider how ridiculous it would be if someone said that the amount of pleasure he has in smoking a good cigar is such that, if it were multiplied by 127, or say by 1077, it would be precisely equal to the amount of pleasure he has in listening to a symphony of Beethoven...”.⁹ Brentano is right, it seems counterintuitive to suggest pleasure and pain are divisible the same way a “foot is divisible into twelve inches”.¹⁰

Feldman provides three constructions of Brentano’s counterexample; the latter being concerned with measuring pleasure.¹¹

1. Some pairs of episodes of pleasure are incomparable in size: there is no number, n , such that one episode is exactly n times bigger than the other.
2. If (1), then [(H)] is false.
3. Therefore [(H)] is false.¹²

⁵ Mill, *Utilitarianism*, 211.

⁶ Leo Zaibert “Just Organic Wholes”, in *The Theory and Practice of Ontology*, ed. Leo Zaibert, (London: Palgrave MacMillan, 2016) 138.

⁷ Ibid.

⁸ Feldman, *Pleasure and the Good Life*, 45.

⁹ Brentano, *Origin of Our Knowledge of Right and Wrong*, 30-1, quoted in *ibid.*

¹⁰ Ibid.

¹¹ Feldman, 48.

¹² Feldman, 48.

This argument seems intuitive, and its implications are damaging for (H). If we accept (1), we concede different pleasures are too fundamentally different to place concrete value on. If we accept “the impossibility of putting a concrete value” on pleasure and pain,¹³ we cannot, as Mill suggests, comparatively rank our pleasures, to create a view of which pleasures are most conducive to a welfare maximising life.

A criticism of this view is that the concrete measurement of pleasure and pain is not required for the ranking hedonism demands. Binder, for instance, argues the objection doesn’t damage the claim that “the relevant welfare concept should... aim at pleasure”.¹⁴ Feldman, too, argues that hedonism only requires the episodes of pleasure and pain arbitrarily “have some size”.¹⁵ This is reasonable, we do not need a concrete and precise measurement system for pleasure to conclude that the pleasure I receive from a bottle of nauseating white wine is less than the pleasure I receive from a bottle of my favourite red. It seems reasonable to concede that concrete measurement is not required for these common-sense type evaluations of pleasure. In the scheme of one’s life, however, this is insignificant. A successful theory of wellbeing demands we are able to evaluate more than small pleasures and make complex choices about where we allocate large portions of time such as to our family, children, partners, and in our working lives. As demonstrated in the first counterexample, many of these things are simply incomparable under hedonism. If hedonism is to be a successful theory of wellbeing, it must provide a way to rank important, nuanced and fundamentally different episodes of pain and pleasure. A more sophisticated axiology accounting for these fundamental differences is required to overcome the reductive ‘homogenisation’ that characterises (H).

The above has argued that through the immeasurability of pleasure and pain, hedonism fails as a successful theory of wellbeing. It has described a theory of hedonism, Brentano’s critique drawn from Brentano’s cigar counterexample, and the subsequent argument formulated by Feldman. Failings established, it discusses one objection to this critique, and offers a potential response in defence of the objection, concluding that hedonism fails as a successful theory of welfare.

¹³ Martin Binder, *Elements of an Evolutionary Theory of Welfare: Assessing Welfare when Preferences Change*, (Abingdon, Oxford: Routledge, 2010) 136.

¹⁴ *Ibid*, 137.

¹⁵ Feldman, 49.

Bibliography – Chicago

- Binder, Martin. *Elements of an Evolutionary Theory of Welfare: Assessing Welfare when Preferences Change*. Abingdon, Oxford: Routledge, 2010.
- Feldman, Fred. *Pleasure and the Good Life*, Oxford: Oxford University Press, 2004.
- Mill, John Stuart. "Utilitarianism". In *The Collected Works of John Stuart Mill, Volume X – Essays on Ethics, Religion and Society*, edited by John M. Robson. Toronto: University of Toronto Press, London: Routledge, 1861.
- Rubin, Michael. "Subjective Theories of Value: Hedonism." Lecture, University of Western Australia, Crawley, WA, March 9, 2020.
- Zaibert, Leo. "Just Organic Wholes". In *The Theory and Practice of Ontology*, edited by Leo Zaibert, 138. London: Palgrave MacMillan, 2016.

PHIL3005: Continental Philosophy

Jack Meakins

Essay Question: Give a summary and critique of Martin Heidegger's 'The Thing'.

Summary

Upon a first reading of 'The Thing', one might be forgiven for feeling deceived; that in spite of his attempts to elucidate some concept of 'nearness', Heidegger, in his unduly vague use of the word, had proved only to further distance one from its true meaning. While it would seem that technologies, such as the radio, do indeed impart some sense of communicative closeness, Heidegger contends that this is not what it means to be near – a vast departure from conventional use of the word. Thus, it could be understood that in asserting as such, Heidegger is implicitly revealing the aim of his work; to illustrate his concept of 'nearness' – via a descriptive inquiry into the encountering of a jug.

To do this, Heidegger makes what he believes to be a necessary distinction between two ways one might view a jug. Firstly, one may experience a jug *qua* jug – a jug in its capacity as our conceptual, and perhaps scientific, understanding of what a jug is¹. A jug is created, necessitating a creator, with a certain material, a base and sides, and with a void. While these aspects are certainly constitutive of how we experience a jug, Heidegger contends this to be somewhat primitive. In viewing the jug *qua* jug, one is necessarily conditioned by the scientific perspective, and thus only sees the jug in terms of what said perspective is made to find – an object; a mode of viewing Heidegger finds prohibitive and misconceived².

In contrast to this, in rejecting the scientific perspective, one may view the jug as a product of its 'jugness' – constitutive aspects or associations one may have when encountering a jug as a 'thing' vis-à-vis an 'object'. Not only is a jug made for the purpose of pouring, but for pouring *out* – perhaps at a feast as means of consecration. In fact, Heidegger has certain associations in mind – what he would deem the 'fourfold'³. To be a thing, the jug – as a product of its 'jugness' – must be encountered as a means to unite the earth, sky, mortals and the gods. Heidegger contends that, given these associations are of little importance to the scientific perspective, they would be necessarily overlooked.

¹ p 166 (Page references provided for M. Heidegger, *Poetry, Language, Thought*)

² p 168

³ p 171

With this, Heidegger believes he is equipped with the means to answer his initial question; what is nearness? Heidegger believes that, in discerning the jug as a 'thing' rather than a mere object, one has gone through an active process of 'nearing', insofar as one has bridged some conceptual gap necessary for 'Being'. While Heidegger's understanding of Being is not espoused within 'The Thing', it is understood that encountering things as things, rather than objects, is requisite for the manifestation of Being⁴. Thus, nearness, to Heidegger, is to experience a thing as a thing – a product of its 'thingness'⁵. It may be necessary to point out that this may be the source of Heidegger's seeming contradiction, as its somewhat obscure definition is hardly congruent with its conventional use.

Critique

One aspect of 'The Thing' which may stand out to readers is Heidegger's argumentative style. Orthodox methods of rational justification are seemingly foregone for poetic and dramatic devices, with an emphasis given to illustrating his conclusions rather than drawing from reasoned premises. On the surface, this may seem a likely target for critique; how could Heidegger reasonably defend his claims from the analytic predispositions of Western metaphysics? However, it would only take a glance at his work to ascertain how he would respond. While an interpretation of 'The Thing' as Heidegger's attempt to illustrate 'nearness' is valid, it could also be viewed as somewhat of a diatribe against the scientific perspective and its various progeny. With the concepts of logic and reasoning being imbedded as institutions of the scientific perspective, it is fairly obvious that Heidegger would dismiss their use as examples of the fundamental restrictions on encountering things as things, as opposed to mere objects. It is thus understandable that Heidegger, in asserting this, would necessarily avoid using such an argumentative framework. While this by no means implies he is correct in doing so, it at least provides a rationale for why he does. Hence, as a preface, I will avoid using Heidegger's argumentative style as a means of critique, on the basis of the principle of charity. However, this is not to say he does not assert premises – perhaps implicitly – that seem dubious at best.

I first want to consider two key assertions made. That 1) The scientific perspective – viewing things as objects, reduced down to their form and function – necessarily conditions one to only look for truths to which the perspective pertains; and 2) Heidegger's perspective – viewing things as things – is determined by a thing's means of uniting the fourfold of earth, sky, mortals and the gods. With respect to his second premise, is it not fair to say that revolving 'thingness' around this fourfold

⁴ See Martin Heidegger's 'Being and Time'

⁵ p 179

ontology be just as reductive? If Heidegger is asserting that the scientific perspective is wrong in only looking for truths to which it permits, could it not also be said that understanding a thing's 'thingness' in respect to a finite number of associations (earth, sky, mortals and the gods.) be just as misconceived? It doesn't seem at all clear to me how these premises don't necessarily contradict each other.

A response Heidegger may have to this is that this fourfold is the **only** respect in which a thing can exist as a thing, and that it thus not restrictive of the truths insofar as it, as an absolute ontology, permits all possible truths. However, at least in 'The Thing', Heidegger shares no reasons as to why it are these four aspects **specifically** that constitute this, merely stating they do *ipso facto*.

Secondly, I would like to assess Heidegger's use of the jug. Heidegger's main objective in illustrating his work is to show that viewing a thing as an object is comparatively reductive to viewing a thing in virtue of its 'thingness'. While there is certainly a gap between the scientific perspective of the jug – somewhat mundane in form and function – and the associations Heidegger's perspective imparts on it, would it not be fair to think there are other things of which a scientific perspective wouldn't reduce? For instance, let's forego Heidegger's jug for the human brain. The human brain is undoubtedly a thing of which the scientific perspective is concerned. Given the complexity of the form and function of the brain and its processes, could it not be argued that a scientific perspective would in fact offer more nuance and meaning to such a thing, especially in comparison to Heidegger's somewhat theistic ontology? It seems that in using the jug as his only illustration, Heidegger has artificially misconstrued the perceived reduction in viewing things as objects vis-à-vis things.

In conclusion, while Heidegger's ability to illustrate a sense of primitivity concerning the scientific perspective is undoubtedly insightful, 'The Thing' is too reliant on appeals to his fourfold ontology to be convincing.

Bibliography

Heidegger, M. (2001). *Poetry, Language, Thought* (1st Perennial Classics ed.). Perennial Classics.

Heidegger, M., Stambaugh, J., & Schmidt, D. (2010). *Being and Time*. State University of New York Press.

Kaz's Comment:

Well summarised. The critique is also good. Well reasoned re his style and you argue against his ontological stance well — and the human brain is an interesting example, although a little more could have been made of it. Whilst I understand the restrictions imposed by the word limit, some repetition could have been eliminated to allow an extra sentence or two. But honestly, I am nitpicking. Overall it is an excellent submission. Congrats!

UWA Undergraduate
Philosophy Society

uphilsoc.org

© 2021