

# A P R I O R I

The Brown Journal of Philosophy

Volume 1



Brown University  
Department of Philosophy  
2016

Editors:

Jacqueline Dewart (Ethics and Aesthetics)

Gregory Heeren (Philosophy of Mind, Language, Science, and  
Logic)

Tin Oreskovic (Metaphysics and Epistemology)

Polina Sandler (History of Philosophy)

Margot Witte (Ethics and Aesthetics)

Head of Design: Polina Sandler

Head of Production: Tin Oreskovic

Managing Editor: Mariana Castro

Editor-in-Chief: Benjamin Seymour

Printed by IngramSpark

[a-priori@brown.edu](mailto:a-priori@brown.edu)

Copyright © 2016 by Benjamin Seymour

All rights reserved. This journal or any portion thereof may not be reproduced or used  
in any manner whatsoever without the express written permission of the  
editor-in-chief.

A Priori: the Brown Journal of Philosophy is made possible through the generosity of the Department of Philosophy at Brown University. The Journal would like to offer special thanks to Professor Bernard Reginster, Professor Richard Heck, and Katherine Scanga for their indispensable contributions.

## Letter from the Editor

A Priori: the Brown Journal of Philosophy is Brown University's first philosophy journal and is devoted to showcasing the finest essays written by undergraduates both within the Brown community and beyond. In our first edition, we the editors of A Priori have sought to combine these academically rigorous essays in such a way as to highlight the depths and contradictions of philosophical thought itself. Covering topics in science, politics, history, and theology, these essays embody the intellectual diversity that is and has been integral to the tradition of philosophy. Furthermore, the inclusion of essays that examine the same issue from opposing perspectives reveals the uncertainty that forever underlies even the most lucid examples of reasoning and the sturdiest edifices of human thought. To read these essays then is to participate in the most important philosophical activities—to critically engage with a text; exercise discerning judgment; and question one's own beliefs. Through confronting the ideas of others, we ultimately illuminate ourselves.

Benjamin T. Seymour, Editor-in Chief

## Contents

A Case for the Guilty: A Defence of Blameless Guilt .....	1
<i>Samantha Wesch, University of Alberta</i>	
A Comparative Analysis of Bonaventure's and Aquinas' Epistemologies .....	26
<i>Luke Wajrowski, University of Notre Dame</i>	
Fairness, Free-riding, and the Distribution of Meaningful Work: Is Guaranteed Employment Superior to Guaranteed Income? .....	40
<i>Lloyd Soh, Brown University</i>	
Gettier vs. Justified True Belief .....	63
<i>Dominic Sicilian, University of Miami</i>	
Am I Doing Bad Science? .....	77
<i>Margaret Min, Brown University</i>	
Pornography as Group Libel .....	95
<i>Irmak Aydemir, McGill University</i>	

Intuitions in Philosophical Thought Experiments ...120

*Abel Girma, Brown University*

Pleasure and the Nature of Complete Acitivity .....134

*Cait Duggan, The University of Chicago*

Eternalism and Presentism .....152

*Jacob Beck, Brown University*

## **A Case for the Guilty: A Defence of Blameless Guilt**

Samantha Wesch

**Abstract:** By uncoupling guilt and blame through a revised fittingness account of guilt, this paper suggests the Kantian account of judgements of blameworthiness (as only applying to what is in the agent's direct control) may be preserved without being incompatible with intuitions regarding the fittingness of guilt towards harms one is merely causally connected to. This paper will take up the fittingness account proposed by Allan Gibbard, and revise the understanding of guilt's proper object and the judgement it elicits in order to understand guilt as meshing with other moral emotions in addition to guilt. I suggest the proper object of guilt is genuine threats to personal relationships, and guilt acts to elicit the judgement of sympathy and caring for others.

**Acknowledgements:** I would like to thank Dr. Howard Nye and Dr. Alan McLuckie for their guidance, support, and kindness throughout this project.

Imagine you receive a call from a friend; while driving carefully and responsibly, she struck and killed a child who darted out in front of her vehicle. The police and witnesses agree there was nothing she could have done to prevent the accident, that this was not a case of even minor negligence, and the accident was completely out of her direct control.<sup>1</sup> When you arrive at your friend's home you find her wracked with guilt. To comfort her, you remind your friend the accident was not her fault as there was nothing she could have done to prevent it, and, since she is not to blame, she should not feel guilty. Having not considered blame as necessary for guilt, your friend perks up is relieved of all her guilty feelings. She no longer feels a sense of personal responsibility or special relationship to the child or the accident. Or, imagine your friend has already come to this conclusion on her own by the time you arrive, free of guilt she proclaims her mere causal role is unfit for blame.

To think one's friend would feel no guilt, despite having no direct control, about her causal role in the child's death is unsettling, if not disturbing, but this reaction does follow from Kantian blame and blame as necessary for guilt's fittingness. The Kantian holds an agent may only be blameworthy for the orientation of her will; therefore, as long as she has acted in good will the agent cannot be blameworthy, regardless of the consequences of the act (therefore, morality exists

within actions themselves, as opposed to evaluations of the outcome of actions):

A good will is not good because of what it performs or effects, not by its aptness for the attainment of some proposed end, but simply by virtue of the volition, that is, it is good in itself, and considered by itself is to be esteemed much higher than all that can be brought about by it in favour of any inclination, nay, even of the sum total of inclinations. Even if it should happen that, owing to a special disfavour of fortune...this will should wholly lack power to accomplish its purpose, of with its great efforts it should yet achieve nothing, and there should remain only the good will (not, to be sure, a mere wish, but the summoning of all means in our power), then, like a jewel, it would still shine by its own light, as a thing which has its whole value in itself.<sup>2</sup>

That the friend did everything she could to prevent the accident (was a vigilant driver, had her car serviced regularly, etc.), not merely that she did not intend to cause harm, leads the Kantian to conclude the friend is not blameworthy. It follows, if blameworthiness is the proper object of guilt, she ought not feel guilty.<sup>3</sup> But this is highly counterintuitive; there is something unsavoury about the friend's guiltlessness towards the child's death she is causally connected to. Further, this intuition seems grounded in more than mere social norms; the friend appears to fail to recognize something morally important about her role in the

child's death, despite her lack of blameworthiness. The discomfort brought by the absence of guilt cannot be explained away; though she is not blameworthy, intuition suggests the friend has some sort of moral responsibility to the child's death brought about by her causal role, and this appears to warrant guilt.

This sticky situation leaves the Kantian with three options: 1) intuitions about guilt and moral luck are mistake; the agent is right not to feel guilty, 2) there must be revisions to the Kantian account of blame to accommodate feelings of guilt when one has merely a causal role in harms (she must accept it is possible for an agent to be blameworthy for what is beyond her control), or 3) the Kantian should uncouple blameworthiness and guilt's fittingness as to allow for situations in which guilt is fitting for an agent who is not blameworthy, and understand blameworthiness as a sufficient, but not necessary condition for guilt to be fitting. Thomas Nagel argues Kant was wrong to deny the existence of moral luck (that an agent may be judged as praiseworthy or blameworthy for events out of her control);<sup>4</sup> and that judgments of praiseworthiness and blameworthiness must, at least somewhat, relate to outcomes of an agent's actions, as it is consequences which are the typical object of moral evaluation:

Prior to reflection it is intuitively plausible that people cannot be morally assessed for what is not their fault, or for what is due to factors beyond their control...However jewel-

like the good will may be in its own right, there is a morally significant difference between rescuing someone from a burning building and dropping him from the twelfth story window while trying to rescue him.<sup>5</sup>

This paper will suggest the Kantian does not need to revise her understandings of blameworthiness in order to account for responsibility to harms one has played a causal role, and maintains the traditional Kantian account of blame. I argue she should accept the third option above, and revise her understanding of blameworthiness as necessary for guilt's fittingness in order to preserve the Kantian account of blame without compromising the fittingness of guilt in causal roles played in harms. This paper argues the judgement elicited by feelings of guilt are that genuine threats to the agent's interpersonal relationships are present and sincere sympathy and regret for harms caused to others, and functions to motivate reparation of damage to social bonds done by the harms to which one is causally connected. Therefore, the proper object of guilt is not one's own blameworthiness (as most fittingness account of guilt suggest), but it is a special obligation to amend the harms which one has caused, regardless of the agent's blameworthiness for the event. The Kantian should then recognize an agent ought to feel guilty for her role in the harms she is causally connected to, but not blameworthy for, as duties of social obligation

create a moral responsibility to repair and reconcile these harms and express genuine care.

Though there is some dispute as to the exact definition and criteria for fittingness, for simplicity I will use a more generic understanding.<sup>6</sup> To judge an emotion as a fitting response is to hold the proper object of the emotion is present, and the judgement the emotion functions to elicit is accurate.<sup>7</sup> Fittingness evaluations are primarily concerned with the rationality of emotions; fittingness examines how one ought to feel, and why. The Kantian understands morality and reason as inseparable, for what it is to be moral is to act with reason,<sup>8</sup> and so understanding the rationality of moral emotions is an important issue for the Kantian.<sup>9</sup> To be fitting, the emotion must fulfil its purpose; they are reasons for why one ought to feel in certain ways, and offer an explanation of the relationship between reasons and values.<sup>10</sup> For example, losing a favourite pair of mittens, or a close friend moving away are scenarios when sadness is fitting. If the proper object of sadness is the loss of something which is valued by the agent (the mittens, the friend's proximity) then sadness' proper object is present in both scenarios. If the judgement elicited by sadness is a valued object is lost, this also occurs in both scenarios. The presence of the proper object and the appropriateness of the judgement makes sadness fitting in both cases. The proper object and judgement are distinct but clearly connected; the emotion must appropriately correspond to the reality

of the situation (the presence of the proper object) and the agent must herself react properly (the judgement elicited). Proportionality is also a factor; that which is of more value to the agent should cause stronger and longer lasting sadness; if her friend's proximity is valued more by the agent than the mittens, it would be fitting for the agent to be more sad regarding the move than the mittens. Fittingness also evaluates when one ought not feel an emotion; sadness is unfitting when the garbage man takes away the trash because, though there an object is lost, it is not a loss of a something valued by the agent nor would the judgement an object of value is lost be appropriate. I will regard an emotion as fitting if and only if the proper object is present and the judgement elicited by the emotion accurately corresponds to the situation.

The origin of emotions may lie in their adaptive value, but this does not compromise the importance of their contemporary social and moral role as eliciting value judgements. D'Arms and Jacobson suggest rational sentimentalism to explain the interaction between the evolutionary history of emotions and their contemporary functions. Rational sentimentalism refers to the idea that emotions correspond to "human values", which are held by every morally regular human and serve as the basis for morality (these values are, however, still subject to reason). Human emotional tendencies do not need to be assumed fitting, fittingness evaluations do need to correspond to actual human

concerns.<sup>11</sup> D’Arms and Jacobson argue “almost all humans have the emotional capacities necessary for being sensitive to ... human values”.<sup>12</sup> Fittingness theories function to account for what judgement, corresponding to a deeply held “human value”, is being elicited when one experiences an emotion. These human values play important roles in human contemporary social and moral life; they correspond to what is important to agent, and serve as the basis for moral action and care. Fittingness evaluations of emotions should be ground in these universal human values. The intuitions regarding guilt for harms one has merely a causal connection to should not be dismissed because they alert us to a abnormal reaction which does not express the human value connected to the situation. Concerns regarding relationships and values, such as needs of belonging and social acceptance, and of genuine care for the wellbeing of others are examples of human values. Guilt should correspond to a near-universal human survival problem, but this alone does not determine fittingness; reason and contemporary function are also major factors. The rough-and-ready definition of guilt is the admittance to wrongdoings and the acceptance of moral blame, and the uncomfortable sensations associated. Cases such as the one discussed earlier reveal the social and moral function of guilt to be much more complex than this common conception. Most fittingness theorists hold it is fitting to feel guilt if and only if one is blameworthy. Allan Gibbard’s account

relies on this relationship between guilt and blame so much so that he understands blameworthiness in terms of guilt’s fittingness.<sup>13</sup> Gibbard himself could hold it be fitting for an agent to feel guilt for events out of her control (as he does not subscribe to the Kantian account of blame), but this would necessarily entail she is blameworthy for the event. To preserve Kantian blameworthiness in Gibbard’s account, one would have to counterintuitively concede guilt is only fitting for harms one is in direct control of; a stark deviation from the traditional Kantian account.

Contemporary evolutionary psychology argues guilt evolved as a response to fear of social rejection when an agent has been unfair to and/or harmed others,<sup>14</sup> but intuitions about cases of moral luck reveal guilt has a much more complex function and evolutionary history than these more shallow accounts. If emotions are understood to be the manifestation of judgements regarding human values, rather than simply evolved behavioural mechanisms to help ensure survival, emotions must be understood both through their evolutionary history and to what “human value” they correspond to, and that these are deeply intertwined. D’Arms and Jacobson’s account meshes well with common attitudes and intuitions about emotion: “people do not care simply about whether emotions do them good, or are useful to nicety at large; we also care about having fitting feelings and the values they decry”.<sup>15</sup> It is not solely about what emotions do for

us (as argued by evolutionary psychology accounts), but also what they mean to us. We feel sad when a friend moves away because it shows we recognize that something of value, the friend's proximity, is lost; even though sadness is unpleasant and temporarily impacts functioning, it would be odd to desire one's friend or oneself not to feel sad about the move, or that they not be consoled by the fact they may call or visit. In our original case, we do not want the friend to feel guilty because of what guilt does (it appears to do nothing but make her feel rotten), but we want her to feel guilty because that means she holds the shared human values and properly recognizes and responds to them (as it represents that she values the child's life and acknowledges her role in the accident). Her lack of guilt suggests she does not hold or cannot recognize and respond to these human values, and indicates a problem in moral functioning. Intuitions about the judgement guilt elicits suggest one should feel guilt for any harms she is attached to, despite blameworthiness. Therefore, the judgement corresponding to guilt is not that one is blameworthy, but must correspond to harms one has played a role in more generally.

I purpose to combine rational sentimentalism with the structure of adaptive fittingness proposed by Allan Gibbard. Gibbard suggests a fittingness account of guilt which explains guilty feelings as originally arising as a socially adaptive reaction towards other's anger when one has acted in a blameworthy fashion.

Gibbard understands emotions as having evolved to coordinate social interactions; they are meant to both convey a judgement one has made (when combined with rational sentimentalism, as corresponding to a particular human value) and to generate a response (when combined with rational sentimentalism, a reaction corresponding to a judgement regarding a specific human value) in another. Gibbard argues emotions must be directed in order to be fitting; in the mittens case one's sadness is directed towards the (absence of) mittens. One is not sad globally (though experience of sadness may be overwhelming), but sad locally about the lost mittens; the feeling is eliciting a particular judgement towards a particular object. Further, since their evolved function is coordination of actions and expectations through physical expressions (for our purposes, of shared human values), emotional expressions should create responses in others<sup>16</sup> (for our purposes, also corresponding to shared human values). In the case of the friend's move, sadness is direct at the loss of his proximity, and therefore our sadness is meant to cause, likely, both sadness in him but also comfort; he witnesses the expression of how we value his proximity (the human value of the intrinsic worth of interpersonal relationships). Gibbard explains there are both first- and third-person fittingness, as emotions are meant to mesh with one another. In the mittens case, the first-person response is sadness, and the third-person counterpart is sympathy. Therefore

sadness and sympathy mesh; sympathy works to simultaneously try to relieve the sadness of another while expressing a judgement regarding the relevant human value. Emotions were not meant to exist or function in isolation; emotions necessarily fit into a complex coordination of social communication based in human values and group cohesion. Therefore, the evolutionary and moral aspects of the emotion are necessary to fittingness theories; each plays a role in the appropriateness and function of emotions.

Feelings of guilt, therefore, must serve the purpose of simultaneously expressing a judgement regarding a human value, respond to its proper object, and work to invoke a response in another. Gibbard argues “guilt meshes with anger in a special way”;<sup>17</sup> impartial anger, responding to a blameworthy action and eliciting a relevant judgement brings about the reaction of guilt in another, which works to counter the anger of the other and hopefully reestablish social cohesion through expressing a human value relevant to care for the unjustly harmed other. Gibbard holds the proper object of guilt is one’s own blameworthiness, as it will counter impartial anger as a third-person response to the blameworthy harmful action. Therefore, guilt and anger both respond to blameworthy actions, guilt is the fitting first-person reaction, while anger is the fitting third-person reaction, and they evolved to coordinate together in a specific way to bring about reactions in the other and solve social problems. Gibbard explains:

Guilt [is the] first-person counterpart to anger ... Not that guilt is self-directed anger; feeling guilty is different from feeling you could kick yourself. Rather, guilt is coordinated with anger in a special way: it aims to placate anger, and it is governed by the same norms as govern anger.<sup>18</sup>

Gibbard holds anger need not evoke guilt in others (they could react with anger themselves, or fear) in order to be fitting, but these other reactions, will not work to counter guilt and reestablish social cohesion in the same way guilt does (for our purposes, as they convey other values), which Gibbard understands to placate anger directly through motivating attempts to make amends and obtain repentence from the harmed part: expressions of guilt work as “smoothing things out and correcting for partiality”.<sup>19</sup> Therefore, guilt works reestablish one’s belonging in the group, and regain the trust and affection of one’s fellows; its purpose is to counteract impartial anger.

However, if guilt conveys the human value relevant to genuine sympathy and regret at causing harms and functions to repair threaten relationships, guilt’s proper object cannot be one’s blameworthiness. Not every situation in which genuine threats to relationships are present is impartial anger fitting on behalf of the third-party (such as the case with the friend’s accident), or the agent blameworthy (in the Kantian sense). Though her actions do not warrant anger, as she is not blameworthy, they certainly will affect her

social standing; she has caused harm to another, and therefore other's trust and affection for her is genuinely threatened. Other's feeling hurt, scared, or unable to trust because of harms one has experienced to which another is causally connected to are all reasons which may be devoid of anger but still threaten one's social relationships. The ability of emotion to aid in moral judgements based on intuitions, in addition to the importance placed on emotion in circumstances where one has played a causal role in harms to others is what makes fittingness evaluations particularly relevant to problems of moral luck and blame. I hold that it is possible to maintain that agents are not blameworthy for harms they play mere causal roles in, but playing any role (including simply a causal role) in harms are circumstances in which it is fitting to feel guilt. For this to be plausible, the blameworthiness must be a sufficient, but not necessary condition for guilt's being fitting. One may worry about how blameworthiness could be understood besides being that which that which it is fitting to feel guilt for. In "Moral Feelings and Moral Concepts" Gibbard argues that, in societies which do not have guilt in their emotional repertoires, shame is fitting in both circumstances where one has acted in a morally blameworthy fashion, and in cases of personal failure; the proper object of shame is present and its judgement is appropriate under both of these circumstances, though morally they are very different. Therefore, shame plays multiple social roles

in these guilt-devoid societies, while maintaining the same proper object and judgement.<sup>20</sup> I propose we ought to understand guilt in a way very similar to this; therefore, Gibbard is right to propose guilt has a special relationship with anger regarding one's own blameworthiness, but this is not the only occasion where it is fitting for one to feel guilt. Therefore, guilt may mesh with various emotions which respond to harms, including one's brought about by harms one is causally connected to, when personal relationships are threatened by events one is connected to. Of course, other's do not experience these emotions for guilt to be fitting, but this is the proper role and function of guilt. We should understand impartial anger and blameworthiness then as sufficient conditions for guilt's blameworthiness, rather than necessary conditions as Gibbard does.

Though Gibbard's account may be able to account for the fittingness of guilt in cases such as our original accident case, it cannot do so without admitting the agent is blameworthy for the harmful event which was out of their causal control.<sup>21</sup> This is incompatible with the Kantian account of blame, and in order to preserve both this account of blameworthiness and intuitions about guilt in cases of moral blame I suggest an account very close to that of Gibbard's, but which uncouples guilt and blameworthiness and slightly revises both the judgement guilt elicits and its proper object, in a similar move that he makes in

regards to the uncoupling of shame and impartial anger in society's without the presence of guilt. In this revised account, guilt functions to alert agents to events which they are causally attached, regardless of the agent's blameworthiness. Causing harm to another will threaten and compromise social relationships and group cohesion whether or not one is blameworthy for the harms done, so guilt should function in the same ways whether or not one is at fault for harms. Guilt therefore meshes with feelings of pain one is causally attached to, rather than impartial anger as Gibbard suggests. Guilt, therefore, motivates not actions which merely counter aggression but those which seek to repair the actual harm caused; guilty feelings respond to the actual harms the agent has a causal connection to, rather than the emotional reactions to the harms, as in Gibbard's account. This account coordinates emotions with one another as well as the event itself; as opposed to merely countering other's reactions, it serves to also express a judgement about the harm. The judgement elicited by guilt then corresponds to the genuine value of other's wellbeing. This is much closer to how guilt's contemporary function; we should think the friend in the original case should feel guilty because she values the child's wellbeing, and is acknowledging her role in the harms to the child in the appropriate way. The proper object of guilt is therefore not when one is blameworthy, but when there are special obligations to make amends for harms one has caused

(either deliberately or accidentally), regardless of any particular relationship. This moves the focus away from preserving social cohesion through countering anger, and onto holding guilt as eliciting the judgement that the wellbeing of others is valuable; this of course functions to show that we do genuinely care about others, and fosters trust, which is integral for social functioning and group cohesion (therefore, guilt is fitting in cases where one causes harms even when do not have a relationship with the harmed; guilt would be fitting if one were to shoot strangers on a desert island from a plane, for example). There is, therefore, a moral responsibility connected to harms one has caused that is distinct from blameworthiness for the harm.

One may wonder how an agent can have a moral responsibility to reconcile harms they are attached to, without admitting the agent is blameworthy. Jeff McMahan offers distinctions between kinds of "attackers" to account for different kinds of causal roles an agent may play in harmful events they are connected to. McMahan defines "culpable attacker" (or "culpable threat") as an individual who is blameworthy for an unjustified threat they pose to the wellbeing of others, and "innocent attacker" (or "innocent threat") as one who is not blameworthy for an unjustified threat they pose.<sup>22</sup> McMahan provides a further subgroup of the "innocent attacker", which he refers to as the "inadvertent attacker" who poses a threat to

the wellbeing of others, which is both unforeseen, accidental, and not due to the agent's recklessness or negligence; the "inadvertent attacker" is the victim of pure moral luck.<sup>23</sup> McMahan uses the term "attacker" to indicate the agent is morally responsible; therefore blameworthiness for harms is not a necessary condition for having responsibilities connected to harms they are causally connected to in McMahan's account. Even if the harm one has caused is completely out of their control, their role in the harm creates a special responsibility to the situation: "[E]ven if the reason that the Attacker or Threat poses a threat is simply that he has bad luck, it is still true he is morally responsible".<sup>24</sup> The fittingness account I suggest has this moral responsibility to harms one is causally connected to, independent from one's blameworthiness, as the proper object of guilt, rather than blameworthiness as in Gibbard's account. If this moral responsibility, which I suggest (in contrast to McMahan) is a special obligation to make amends for the harms one has caused, is present then it is fitting for the agent to feel guilty.

One may be concerned this is too weak an account of fittingness; that is, if the proper object of guilt is a special obligation to make amends for harms one is causally connected to, and the judgement feeling guilt elicits is genuine sympathy for harms one is causally connected to, one may worry this account allows guilt to be fitting in situations where it should not be. Imagine a professor finishing up final grades for the semester.

One of the professor's students, who was an attentive and polite student, but despite her hard work, constant attendance and visits to office hours, the student has performed poorly in the course. This student goes to the professor and asks her to boost her grade. The student explains that if she does not get at least a C in the course she will be put on academic probation, lose her scholarships, and will not be able to finish her degree. For her own moral integrity, and out of fairness to her other students, the professor refuses to increase the grade. Months later, the professor learns the student has lost her funding and dropped out. The fittingness account I have suggested holds it would be fitting for the professor to feel guilty in regards to the student's failure to continue her studies. The proper object of guilt, a causal connection to harms to another, is present, and the judgement elicited by guilt, genuine sympathy for the harms to the other, are both present. For it to be fitting for the professor to feel guilt may appear counter-intuitive; the professor is not at fault for the student's failure, and the professor acted in a morally praiseworthy way (to deny the student an unfair GPA boost). However unfair it may appear, I maintain it is fitting for the professor to feel guilt for the student's drop out. The "fairness" of one feeling the emotion has no bearing on its fittingness; fittingness is determined solely by the presence of emotion's proper object and whether the judged which the emotion represents corresponds to the situation. The professor

has played the role of an “inadvertent attacker”, she has caused the student harm (the student did try her best) but her role in the harm was due to only bad luck. Here is D’Arms and Jacobson:

[T]here is a crucial distinction between the question of whether some emotion is the right way to feel, and whether the emotion gets it right ... [A]n emotion can be fitting despite being wrong (or inexpedient) to feel. In fact, the wrongness of feeling an emotion never, in itself, constitutes a reason for that emotion to be fitting.<sup>25</sup>

Therefore, though we may agree it is “unfair” for the professor to feel guilty when she has done nothing wrong (she has, in fact, done the right thing), guilt is still fitting. Neither does it matter that the professor’s actions were morally praiseworthy, as moral evaluation should not affect the fittingness of guilt. One may hold that, since it is the student’s fault she has failed the course the professor should not feel guilty; the student caused the harm to herself, and therefore the professor’s lesser causal role is unfitting for guilt. This too though does not hold, the presence of a blameworthy agent does not change that the proper object is present for the professor to feel guilt, and that her guilt would elicit an appropriate judgement.

This revised fittingness account of guilt should be accepted by the Kantian so she may perverse the traditional Kantian account of blameworthiness and

reconcile it with intuitions about the fittingness of feelings of guilt to harms in which one was merely causally connected to. This account does not, of course, solve all the problems for the Kantian arising from situations of moral luck (particularly those in which one has attempted to commit a harmful act but is intercepted by mere luck from obtaining their goal) but does allow her a more complex relationship between emotions and moral evaluation. Guilt therefore serves to recognize harms other experiences, rather than an acknowledgement of the agent’s own blameworthiness.

**Works cited:**

1 This is a variation of problems of moral luck as discussed by Thomas Nagel. See: Nagel, Thomas. "Moral Luck" *Mortal Questions*. Cambridge: Cambridge University Press, 1979. pp. 24-38.

2 Kant, Immanuel. "The Fundamental Principles of the Metaphysics of Morals" *Kant's Basic Writing* ed. Allen Wood trans. F. Max Muller and Thomas K. Abbott. New York: The Modern Library, 2001. p. 152.

3 One should not, however, consider the friend's actions necessarily morally praiseworthy, as it is not clear whether her lack of negligence was an act in accordance with duty or for the benefits driving responsibility entails (she needed to get home safely, she could be harmed in an accident, worries about insurance, etc). This is an important distinction for the Kantian, but regardless of her praiseworthiness in acting responsibly, she may not be blameworthy for the accident as she did not act out of bad will. See: Wood, Allan. "The Good Will" *Philosophical Topics*. vol. 31 no. 1/2, *Modern Philosophy*, Spring and Fall 2003. pp. 457-484.

However, it has been suggested that one may have secondary motives in addition to the primary motives of acting from duty while maintaining praiseworthiness. See: Stratton-Lake, Phillip. *Kant, Duty and Moral Worth*. New York: Routledge, 2000. pp. 60-77. Therefore I suggest praiseworthiness is not necessitated by lack of blameworthiness, but this is not integral to our purposes.

4 Nagel, Thomas. "Moral Luck" *Mortal Questions*. Cambridge: Cambridge University Press, 1979. p. 26.

5 *Ibid.* p. 25.

6 Value-fittingness theories are not themselves uncontested. See: Crisp, Rodger. "Review of Value...and What Follows by Joel Kupperman" *Philosophy* 75, 2000. pp. 458-462.

7 Jacobson, Daniel, «Fitting Attitude Theories of Value», *The Stanford Encyclopedia of Philosophy* (Spring 2011 Edition), ed. Edward N. Zalta.

8 For further discussion of reason, morality, and the Categorical Imperative in Kant's account see: Kant, Immanuel, "Critique of Practical Judgement" *Kant's Basic Writing* ed. Allen Wood trans. F. Max Muller and Thomas K. Abbott. New York: The Modern Library, 2001. pp. 223-272.

9 It is of course one has little control of their emotions, however, this does not mean they cannot be judged as rational or irrational. Kant argues an agent ought to take measures to train her responds that are out of her direct control to be in line with morality. This relates to fittingness theories' goal of emotional regulation. See: Kant, Immanuel. "Anthropology from a Pragmatic Point of View" *Anthropology, History, and Education*. ed. Gunter Zoller and Robert B. Loudon. trans. Mary Gregor et al. Cambridge: Cambridge University Press, 2014. p. 347.

Though they may not be in direct control of the will, there are practises one may do to bring about particular emotion dispositions or develop some ability for emotion regulation. See Gibbard, Allan. *Wise Choices, Apt Feelings: A Theory of Normative Judgement*. Cambridge: Harvard University Press, 1990. pp. 38-40.

10 Nye, Howard. "The Wrong Kind of Reasons" *The Routledge Handbook of Metaethics* eds. Tristram McPherson and David Plunkett. New York: Routledge, forthcoming.

11 D'Arms, Justin and Daniel Jacobson. "Anthropomorphic

Constraints on Human Value” Oxford Studies in Metaethics vol. 1 ed. Russ Shafer-Landau. Oxford: Oxford University Press, 2006. p. 116.

12 D’Arms, Justin and Daniel Jacobson. “Anthropomorphic Constraints on Human Value” Oxford Studies in Metaethics vol. 1 ed. Russ Shafer-Landau. Oxford: Oxford University Press, 2006. p. 113.

13 Gibbard, Allan. *Wise Choices, Apt Feelings: A Theory of Normative Judgement*. Cambridge: Harvard University Press, 1990. p. 126.

14 Buss, David. *Evolutionary Psychology: The New Science of the Mind*. Boston: Pearson Education, Inc., 2012. p. 408.

15 Ibid. p. 112.

16 Gibbard, Allan. *Wise Choices, Apt Feelings: A Theory of Normative Judgement*. Cambridge: Harvard University Press, 1990. p. 135.

17 Ibid. p. 140.

18 Ibid. p. 139.

19 Ibid. p. 147.

20 Gibbard, Allan. “Moral Feelings and Moral Concepts” Oxford Studies in Metaethics vol 1. ed. Russ Shafer-Landau. Oxford: Clarendon Press, 2006. pp. 4-5.

21 Ibid. pp. 138, 148.

22 McMahan, Jeff. *The Ethics of Killing: Problems at the Margins of Life*. New York: Oxford University Press, 2002. p. 401.

23 McMahan, Jeff. “Self Defence and the Problem of the Innocent Attacker” *Ethics* vol. 104, January 1994. p. 264.

24 McMahan, Jeff. *The Ethics of Killing: Problems at the Margins of Life*. New York: Oxford University Press, 2002. p. 405.

25 D’Arms, Justin and Daniel Jacobson. “The Moral Fallacy of the “Appropriateness” of Emotions” *Philosophy and Phenomenological Research* vol. LXI no. 1, July 2000. pp. 66, 69.

## A Comparative Analysis of Bonaventure's and Aquinas' Epistemologies

Luke Wajrowski

**Abstract:** Bonaventure's and Aquinas' theories of knowledge are often seen as contradictory; however, I argue that Bonaventure's third-term illumination theory is comparable to Aquinas' two-term theory of knowledge, as the Logos in Bonaventure's illumination model is compressed into the second term in Aquinas's model, forming what Aquinas calls the agent intellect. I will evaluate the validity of this substitution to show that the removal of an explicit third-term is not an effective alternative, for this "substitution" diminishes a direct reliance on the illuminating Logos for knowledge. This paper proceeds from an explanation of the similarities between the two epistemologies to the problems encountered in Aquinas' model and finally explains how Bonaventure's theory of knowledge overcomes these problems by integrating the third-term.

## I. Introduction

Examining the historical context in which Bonaventure developed his illumination theory reveals an important piece of information that is useful in interpreting his philosophy: that he developed his theory at the emergence of the translation of Aristotle into Latin. Aristotle's texts and commentaries had a limited influence on Bonaventure and the development of his philosophy. While Bonaventure hesitated to embrace Aristotle's philosophy, Bonaventure's contemporary, Thomas Aquinas, wholeheartedly adopted Aristotle's wisdom and methodology. Because of these two different approaches to the work of Aristotle, these two philosophers, Bonaventure and Aquinas, are often viewed as having developed opposing philosophies—in particular opposing epistemologies. Bonaventure's epistemology is grounded in his illumination theory, which has its roots in Augustine and neo-Platonic thought; Aquinas' is established in his theory of knowledge involving the agent intellect, developed out of the Aristotelian tradition. However, by taking the three-term model for Bonaventure's illumination theory (involving the human knower, an object of knowledge, and the Logos (the principle of knowledge) as explained by John White in his article *Divine Light and Human Wisdom: Transcendental Elements in Bonaventure's Illumination Theory*, I propose that the two theories aren't at odds with the other, but that the Thomistic notion of the agent intellect accounts for the

third-term Logos in Bonaventure's illumination theory. This comparison, at face value, seems tricky, but the substitution of the third term Logos in Bonaventure's illumination theory for the natural light of the agent intellect in Aquinas' epistemology will be spelled out in what follows along with the implications of making such a "substitution."

I will first delineate Bonaventure's illumination theory, pulling from Bonaventure's own writings as well as more modern takes on Bonaventure to reconstruct a three-term theory that accurately portrays the illumination theory. I will then focus on Aquinas, turning to his notion of the agent intellect and how the natural light of the human knower parallels Bonaventure's model of knowledge, which involves illumination from the Logos. Finally, I will evaluate the validity of this substitution to show that the task at hand is an immensely complex one that requires an all-encompassing understanding of the implications each position makes.

## II. Bonaventure's Illumination Theory

Bonaventure's epistemology is founded upon God, who is the First Truth and the "adequate and actual Cause". The reason for such a foundation involves the fact that the material world—including the human knower is in constant flux; the stability that emanates from God as the first principle of knowledge (Who doesn't change) can condition certainty of knowledge

(knowledge that is infallible and immutable). Thus, the divine, acting as a first cause, conditions certainty in knowledge. Bonaventure asserts a reliance upon the divine in philosophizing primarily because He (Who is Truth Itself) encompasses all truth. In his *Disputed Questions on the Knowledge of Christ*, Bonaventure delineates two positions of obtaining knowledge through the Logos. One view is that the uncreated wisdom manifested in the eternal forms is so sublime that it "can never be attained"; the other extreme is that human knowledge is unchangeable and eternal reason only has a mere "influence". The former relies too heavily on the objective transcendence such that knowledge can never be arrived at through the human intellect; the latter asserts the sublimity of the human person, not taking complete account of ignorance or the limitations of the human perspective. One absolutely transcends human capabilities; the other doesn't allow for much transcendence.

Bonaventure rejects both of these positions and proposes his own between the two extremes, one that realizes the capacity of the human intellect to reflect on the world's intelligibility while respecting that absolute and eternal truth cannot be compressed into a finite understanding. Drawn out to the proposition's logical conclusion, certainty is a "function of both intentional and participatory relationships".

Bonaventure develops his middle course in reaction against the limitations imposed by the two extremes.

The first position in which knowledge is conditioned by eternal reason to the human mind produces skepticism about the changing material world. Bonaventure rejects this outlook because of its intellectual absurdity in dismissing the corporeality and the temporal order, which is innately intelligible. The other extreme is that the divine intelligence only shines forth in the “essences of things” and not in the act of obtaining knowledge about these things. Bonaventure rejects this viewpoint because of the fluctuating relationship between objects and subjects. Material objects are known in instances, because not only do the objects themselves change, but the subject observing the changing objects is also changes. The varying relationship between objects and subjects cannot produce certainty (constancy of knowledge) merely in this manner, but can only produce instance of knowledge, i.e., a certain object known by a certain subject at a certain point in time.

Bonaventure, deeming the two extremes epistemologically unsatisfying, proposes his *via media*, steering between the two erroneous extremes. His middle course, known as his illumination theory, engages not merely two terms (the divine and the subject or the subject and object in the two extreme positions) in developing knowledge but includes the third term necessary to produce certainty. Wisdom attained by a human knower is a participation in the light of the eternal wisdom—which both illuminates the human intellect and the essence of the object of

knowledge. The eternal wisdom is nominally the Logos, the second Person of the Trinity, through Whom, in Whom, and for Whom, all creation was made. The Logos contains the unchanging essences of all of creation. The relationship between the eternal Logos and the temporal created world is predominantly two-fold. In relation with the object (indeed all objects) of knowledge, the Logos shines down the divine ideas upon all of creation, and the object of knowledge receives its essence insofar as it participates in its divine idea. The Logos also illuminates the subject, so that the subject can see the essence of the object in the elucidating hierarchy of being. The subject can clearly perceive the essence intimated in the object in the invisible light of the Logos.

The knowledge acquired in divine illumination participates in the Logos, as the objects of knowledge participate in the divine ideas. Because of the eternal character of knowledge obtained in the divine light, opposed to the mere instances of truth grasped through the relationship between solely an object and subject, the illumination theory produces a certainty known as “created wisdom”. The created wisdom, although certain, is limited by the finitude of the human knower, but is nonetheless a participation in the unconditional, unlimited eternal wisdom associated with the Logos.

It is important to realize that Bonaventure did not see the Logos as completely separate from temporality and merely shedding light upon the subject-object

relationship, but as the necessary bond for knowledge, imparting intelligibility upon all of creation and illuminating the subject to understand not merely a particular object, but a particular object in the whole hierarchy of being. The relationship between the human subject (whose soul by nature is connected with the divine) and the Logos is not characterized by a distant, indirect illumination, but is a connection which cannot be dismissed—as is the relationship between the object and the Logos, who imbues intelligibility into the dust of creation.

### III. Aquinas' Theory of Knowledge

Accepting the three-term model for Bonaventure's illumination theory makes evident its similarities to Aquinas' theory of knowledge. Aquinas' theory of knowledge is a linear process in which a further step follows the completion of its previous step; with one of the steps missing, the process for knowledge about the particular object of knowledge cannot be completed.

Aquinas reasons that all knowledge begins in the senses. The sense faculties, powers of the soul, are the channel through which the objective, outside world enters the subject in order to be known. Although the process for knowledge begins with the sense experience, not all knowledge is sensory knowledge (as will be shown later in the process). After the outside world become available to the intellect through the senses, the object is converted to a phantasm. The

phantasm represents the object by an image in the mind. The characteristics of the specific object picked up through the senses are imprinted in this image of the mind. From this step, the agent intellect illuminates the phantasm, extracting the general essence from the specific image. The agent intellect, Aquinas reasons, is a power of the intellect itself. Each human person has his or her own agent intellect, which, insofar as each intellect functions, is a participation in the divine Intellect. In this sense, the power of the agent intellect, to come to knowledge, resides inside each individual knower.

The essence extracted from the phantasm is impressed upon the possible intellect. The possible intellect receives the essence of the object. The possible intellect then cognizes the essence and understands the nature of the object illuminated. The final step involves a *verbum mentis*, a “word of the mind,” to express the cognition. The concept formed is useful in expressing the idea of the object as well as recalling it.

In the process of knowledge, the active intellect takes “information that is material and particular” and converts it “into something immaterial and universal.” The whole process is instantaneous and not deliberate; Aquinas' theory of knowledge is an expression of the automatic epistemological process of the human mind.

Applying the tripartite model of knowledge deduced from Bonaventure's illumination theory to Aquinas' theory of knowledge makes evident the substitution

I want to highlight of Aquinas' agent intellect for Bonaventure's Logos. It will be important to note that the following substitution I delineate is a fundamental and simple one; the specific details involved as a consequence of making such a switch will be explicitly articulated in the next section of this paper in order to first emphasize the general principle of the substitution, not whether the substitution is an absolutely effective one.

The primary difference between both theories of knowledge is the role the divine takes in the philosophical model, which is made comparable in both cases in the substitution I will delineate. In Bonaventure's illumination theory, the divine light allows for a clear and illuminating ground for realization. Without the powerful luminosity of the divine, the intellect, dark and ignorant, would not come to the illuminated and certain conclusion attributed with knowledge. However, the luminosity of the divine is comparable to the power of the intellect planted in the human person and the possession of intellectual forms that the material world contains in Aquinas' theory. The human person is implanted with the divine gift of an autonomous intellect. The human intellect, supplied entirely by the divine, relies on the divine for its contingent existence and continuity; however, the divinely endowed capacity does not rely on the divine in the process of coming to an understanding. In this sense, then, the radiating light of the Logos of

Bonaventure's theory is replaced with the innate power of the Logos endowed in the human intellect of Aquinas' theory. The illumination factor of the external, third-term Logos is poured into the human intellect in an essentially two-term model.

In Bonaventure's theory, the Logos, the subject, and the object are necessary for the process of philosophizing; neglecting one of the terms results in uncertainty, and thence knowledge is not possible: the Logos provides the light and truth overshadowing both the subject and object—without the overshadowing light of the Logos, the subject cannot “see” the object. The lack of the subject or object clearly cannot allow for knowledge, as the subject is the one to whom knowledge is attributed and the object is the aim of knowledge.

The three terms collectively allow for the attainment of knowledge. Although Aquinas' model of knowledge is essentially a two-term theory (although the object partakes in the divine ideas), Aquinas' theory is nonetheless able to attain knowledge like Bonaventure's. The stability and coherence of Aquinas' predominantly two-term theory (in contrast to the frequent unreliability and flux of the relationship of other two-term models) is made possible by the indirect combination of the Logos (which sheds light upon the intellect in Bonaventure's theory), and the human intellect, thereby producing the same power and ability as is in Bonaventure's model.

In this way, the substitution of the Logos in Bonaventure's three-term theory is in essence a sort of combination of the Logos and the subject to produce a cognitively self-sufficient human intellect in Aquinas' theory of knowledge.

#### **IV. Evaluating the Substitution**

After spelling out the substitution and consequently unifying the two comparable epistemologies, the relative merits of their divergences must be evaluated. However, there is a particular difficulty in evaluating the effectiveness of one epistemology over another because, many times, a combination of positives and negatives characterize each philosophy, making it difficult to compare two entirely different entities under one rubric. In attempting to conclude which philosophy surpasses the other, I will spell out the benefits and setbacks of both Bonaventure's and Aquinas' theories and will make a final conclusion based on their totalities.

Bonaventure's illumination theory asserts the necessity of a transcendent relationship between a knower and an object. The mutability of the knower can't be trusted, and neither can the inconsistencies of the object. The emanation of the unchanging, eternal Logos allows for the knower to "see" and attain certain knowledge, contextualizing the object in the stable order of being.

Aquinas' theory grasps at the imminence of the divine in the natural world. From the divine intellect "forms flow forth into all creatures," allowing for the sovereignty of the human person in cognizing the natural world. However, from a Bonaventurian perspective, the neglecting the direct radiance of the Logos wouldn't guarantee the certainty at which knowledge aims. Without the illumination attributed to the divine, knowledge is a mere science of the natural world at best.

While this last point is seen through the scope of a Bonaventurian philosophy, it is a crucial one to realize. Although the stability associated with the Logos is somewhat accounted for in the firm human intellect of Aquinas' theory, the independence of the human knower is primarily asserted over the continual dependence on the divine light. Certainly the divine gifted the human person with the intellect, but the stress on the continual recollection of the divine is an important one. In Aquinas' model, the cognitively independent agent intellect doesn't necessarily recall the divine in the act of knowledge, and in turn, neglects to contextualize objects in the broader and no less important reality.

Bonaventure's illumination theory surpasses Aquinas' theory of knowledge on this essential point: the illumination theory brings about not only the certainty attributed to the unchanging, eternal, transcendent reality, but the contextualization of knowledge. The

contextualization may appear as a minor addition to knowledge of an object, but it in principle is a vital element that must not be separated from knowledge of an object; the contextualization of the knowledge intimates an ethical framework of valuing goods in their proper order.

Aquinas' theory implies neglecting the immaterial Light that allows for knowledge in the first place, while grasping onto an object in the natural world to know. Knowledge treated as a natural science lacks its inherent companion, the science of ethics. Because of the separation of natural science and ethics implicitly embraced in the philosophy of an intellectually self-sufficient person, the Bonaventure's illumination theory exceeds Aquinas' theory of knowledge with respect to the moral implication behind a cognitively sufficient individual.

However, this remark must be further qualified. Aquinas' implicit expression of the autonomy of the human intellectual to know an intelligible world doesn't straightforwardly "neglect" the divine; it relied on the divine for the complete gift of the intellect and continues to rely on the sustaining of the intellect.

Another necessary qualification is that the sovereignty of the cognitive element of the human knower does not imply the exclusion of the ethical element of the human person. Nor does it imply that the intellectualizing cannot be done in an ethical manner or setting. Aquinas' philosophy must not be

unnecessarily deemed unethical because it does not explicitly and continually rely on the light of the divine (but upon the gift of the divine).

The issue at hand, i.e. of asserting the "better" of the two theories of knowledge after being unified under a common term—namely, the substitution of the Logos for the autonomous human intellect—is an immensely complex one. One must first realize that one epistemology is not absolutely superior to the other, but that each theory has its strengths and weaknesses that must be accounted for through comparison. Depending on what relative aspect of their epistemologies is the focal point, the strength of one over the other will exceed. Therefore, the position of an absolute superiority of either Bonaventure's illumination theory or Aquinas' theory of knowledge to the other cannot be held.

## Fairness, Freeriding and the Distribution of Meaningful Work: Is Guaranteed Employment Superior to Guaranteed Outcome

Lloyd Soh

**Abstract:** This paper compares guaranteed employment with guaranteed income. At first glance, guaranteed employment seems superior to guaranteed income for two reasons. First, employment or work has an intrinsic value that guaranteed income cannot provide. Second, guaranteed income is unfair because it allows for the idle to free-ride on the rest of society. This paper will argue that those two reasons do not hold. In addition, it will evaluate several reasons for and against the significance of the free-rider argument. Ultimately, I will argue that meaningful work is a scarce resource and those who do meaningful work free-ride on the efforts of those who do meaningless work – creating a need for meaningful activity to be redistributed.

Robert Kennedy once said, “I’m not for a guaranteed income. I’m for a guaranteed job”.<sup>1</sup> While guaranteed employment and income both provide the unemployed with a means of income enough to meet their basic needs, there is a clear difference between the two as indicated by the quote above. Guaranteed employment is able to provide the activity of work, usually through the government as the employer of last resort. Under a system of guaranteed employment, employment is not only guaranteed but required for income: income becomes conditional on accepting the work provided by the government. The case for the superiority of guaranteed employment is two-fold: firstly, work has an intrinsic value beyond income alone; secondly, making income conditional on work avoids the free-rider problem. However, this paper will show that even if work has an intrinsic value, guaranteed employment does not provide work of such value. Furthermore, the free-rider problem is an insufficient reason to reject guaranteed income which deals with the unfair distribution of meaningful activity.

A guaranteed income provides the unemployed with the ability to provide for their own needs. Allowing for citizens to have a certain standard of living is in itself a good thing. No individual should be allowed to persist in a state of poverty and the government has an obligation to prevent such conditions. It can be argued that this government obligation remains even if an individual does not wish to accept guaranteed

employment. If income were conditional on accepting government employment, the individual who does not wish to take guaranteed employment will continue in a state of poverty.

As a result of removing the threat of poverty, the worker is also no longer economically coerced into taking a stultifying job. Such jobs are not only meaningless, but often in poor working conditions and low compensation. In the workplace, the threat of unemployment which enabled employers to coerce their workers would lose its effect. A worker who previously had no real alternative to such a stultifying job is now able to quit because a guaranteed income allows him to meet his needs without taking the job. Furthermore, even if the worker does not quit, the fact that he is able to quit drives firms to improve salaries if not working conditions in order to retain labor.

Guaranteed employment is equally able to provide a means to avoid poverty and worker exploitation. Guaranteed employment brings with it income as long as the individual is willing to work. In addition, government-provided employment provides a reasonable alternative to an undesirable job as well, granting workers the freedom to leave their previous job for the government-provided job. As pointed out by Kalecki, “under the regime of full employment, the ‘sack’ would cease to play its role as a disciplinary measure”.<sup>2</sup> The only difference from guaranteed income is that these benefits of dealing with poverty

and worker exploitation are dependent on accepting government employment.

However, there is nothing wrong with having individuals who persist in a state of poverty if they reject the means out of poverty. The government is only obliged to provide the opportunities for welfare rather than actually determining the actual levels of welfare that individuals reach.<sup>3</sup> Even in the case of guaranteed income, an individual could, in the extreme case, dispose of the income he collects by setting it on fire. In a more probable case, the individual might squander all his income on alcohol. In both cases, the individual is in a state of poverty as a foreseeable consequence of his own actions. As long as the opportunity has been presented, the government has fulfilled its obligation. Hence, in the case of guaranteed employment, even though individuals who refuse to take the government-provided employment will persist in poverty, there is nothing unjust because it is a product of the individual's own choice.

Beyond the distribution of income, guaranteed employment deals with the problem of an unequal distribution of work. In an economy where there are fewer jobs than job-seekers, the government needs to at least provide the opportunity for people who want to work to get work. Work provides some intrinsic benefits which the unemployed miss out on and suffer as a result. Work provides a sense of fulfillment in overcoming the resistance of objects.<sup>4</sup> It is in overcoming

obstacles that we develop our capacities and realize our freedom. Furthermore, the lack of employment tends to erode self-esteem. It is easy for the involuntarily unemployed to attribute their failure to find a job to personal incompetence. Merely guaranteeing income is not sufficient for providing these two values. The involuntarily unemployed individual who receives a guaranteed income may continue to feel like a failure and feel even more like a parasite on society.

These intrinsic values are not unique to paid employment. An individual can easily find the same fulfillment in doing volunteer work or pursuing causes that he believes in. In those activities, the individual can still overcome the resistance of objects and do something socially productive, which will prevent a loss of self-esteem. Adding monetary compensation to an activity does not make it any more fulfilling. If the purpose of guaranteeing employment is to supply fulfilling activity, there is no shortage of unpaid fulfilling activity even in the absence of guaranteed employment.

It must be noted that guaranteed employment does not equate to guaranteeing meaningful work. Not all work is fulfilling and even work that allows the individual to overcome obstacles might be done in unhealthy and dehumanizing conditions. Even those who are seemingly in support of a right to work do so only if it is meaningful. Cullen believes that “the quality of the work done is just as important for people’s

well-being as the quantity”.<sup>5</sup> Similarly, Arneson, who argued in favor of guaranteed employment, suggested nine guidelines for jobs to boost self-esteem, including the requirement that “the jobs provided by the state must be manifestly socially beneficial”.<sup>6</sup> It is unlikely that these proponents of guaranteed employment will support the guarantee of meaningless and stultifying work.

Even if we accepted that work had its intrinsic value, it is likely that guaranteed employment will not provide this sort of work. If the unemployed feel a loss of self-esteem because they feel like they have failed in the market system, needing to rely on the employer of last resort does little to ease this feeling. Having a job in itself is not enough to provide self-esteem. According to Elster, jobs provide self-esteem because they meet some form of societal demand. If there is already a demand for jobs that provide self-esteem, there is no need for the government to provide it. But if the government has to provide for it, it means that there is no demand for such a job, which means that it is unlikely that such a job will be able to promote self-esteem.<sup>7</sup> In other words, jobs that are provided with the purpose of promoting self-esteem will fail to do so. Even if governments set a policy such that only meaningful work would be provided, these would likely be capital-intensive rather than labor-intensive. Such a policy would require huge subsidies, which will not only be costly, but also undermine the

individual's self-esteem.<sup>8</sup> In all likelihood, in order to cater to the typically unskilled unemployed, the work will not require a high skill level and will be labor-intensive. Arneson also suggests that "if the state as the employer-of-last-resort designs jobs that require a high skill level, the truly disadvantaged are unlikely to derive much benefit".<sup>9</sup> Yet it is precisely such work that is unlikely to provide much meaning.

Given that government-provided employment will most likely not be meaningful, guaranteed income has several clear advantages over guaranteed employment beyond their overlapping function of providing income to the unemployed. Firstly, guaranteed income removes the need for an individual to spend time on a meaningless job, hence allowing him to have more free time. This includes not only time for leisure, but also time to pursue activities which the individual deems as meaningful. As mentioned above, there are ways to be socially productive even though it might not be paid for. It is likely that these forms of activities will be more meaningful than the government-provided job, precisely because they were chosen by the individual rather than the government. The freed-up time enables the individual to pursue such activities to a greater extent than he would have with a job.

Secondly, under the case of guaranteed employment, economic coercion is not entirely dealt with. An individual might not be economically coerced into taking a stultifying job offered in the book shipping

industry, but he might nonetheless be economically coerced into taking a government-provided job, which might not be meaningful to him. The need to receive an income still compels an individual to take a job even if he would prefer not to and he has no reasonable alternative to taking the government-provided job. It is equally likely for the government as an employer to impose unreasonable demands on its workers knowing that they have no real option of quitting. Even if the government does not impose unreasonable demands, the fact remains that an individual who does not want to work will still have to take a job to feed himself. In contrast, guaranteed income provides income regardless of whether the individual decides to take a job. An individual will not have to do a job that he feels he is not suited for and has the option of only taking work that he deems meaningful.

Overall, a guaranteed income provides the individual with what van Parijs calls real freedom.<sup>10</sup> Real freedom goes beyond protecting pre-existing natural rights and protects the material means to lead lives as they wish. This includes the freedom from the circumstances of poverty and the freedom to say no to an employer without fearing the unemployment that might result. The individual could also have more time to choose the activities he wishes to pursue instead of spending his time working. Even if guaranteed employment provided the material means to determine one's life, it will not provide the time to do so. Especially if an individual

has a strong preference for leisure and would rather not work, guaranteed income enables him to have the real freedom to pursue such a lifestyle. Furthermore, it would otherwise be unfair for someone who is congenitally lazy to have less freedom than someone who prefers to work. People who are averse to work can pursue their conception of good that of idleness while people who are not averse to work can continue to do so. A guaranteed income allows both groups of people to pursue their conception of the good.

The defender of guaranteed employment will probably need to acknowledge these benefits exist, but argue that they are not of sufficient importance relative to the competing considerations. An individual might be made better off by being given more leisure time, but he will likely be made better off if the government gave him a seaside condominium as well.<sup>11</sup> This does not mean that the government should give him a seaside condominium. The point is that just because a provision makes an individual better off does not mean that the government is obliged to grant it, as they are luxuries that will come at a cost to society. As for economic coercion, at least the individual has a real alternative to what may be the worst jobs in society. Even if the individual is economically coerced into taking a government job, at least he is likely to be better off than being economically coerced into taking a job under an employer trying to maximize profits and minimize costs, including the costs of safety measures.

More significantly, these two benefits are outweighed by other greater competing concerns, such as the cost to society and unfairness in the form of the free-rider problem.

Free-riding occurs when one group of individuals is able to live off the efforts of another group of individuals without reciprocating or contributing in any way. Under guaranteed income, the unemployed are able to free-ride on the efforts of the employed. According to Hart, “accepting a benefit creates a liability to contribute to its cost of production”.<sup>12</sup> By not contributing to the cost of production, the free-rider violates a principle of reciprocity and takes unfair advantage of those who do contribute. Perhaps free-riding would be permissible if it came at no cost to those who put in the effort, but providing a guaranteed income to people who would otherwise work does impose costs. Because there are fewer employed people contributing to production and more people drawing upon it in the form of guaranteed income, the size of the economic pie shrinks. This could be said to be unfair to the employed individuals who put in the effort to produce and contribute to society. They are made worse off as their total produce shrinks in order to support the unemployed who contribute nothing to society.

There are several assumptions in trying to use the free-rider problem to show that guaranteed income is unfair. The first assumption is that those who are unemployed contribute nothing to society just because

they do not accept employment. This is not necessarily true. Rather, the unemployed merely refuse to do what the market recognizes as work. It is possible for them to contribute to society through means which are not recognized by the market. This includes unpaid volunteer work, involvement in political activism or even the raising of children. All of these contribute to society even though there is no compensation through the market. Especially if we hold the belief that human beings have a need to engage in productive activity for fulfillment, it is likely that people will still engage in socially productive activity even in the absence of wages. However, this argument is limited as it does not deal with the individuals who decide to contribute nothing to society. No doubt some may choose to contribute in unpaid means, but not all will. Those who contribute nothing will still be free-riding on the efforts of others.

The second assumption behind the free-rider argument is that there is supposed to be a relationship between the amount of contribution and amount of benefits. Benefits and contributions do not tend to correspond proportionally in real life.<sup>13</sup> In society, individuals vary in terms of their contribution but the benefits they receive in the form of non-excludable public goods are generally the same. Individuals could also vary in terms of their benefits: an individual who falls ill more often benefits more from the provision of subsidized healthcare than a healthier individual, even if the two of them contributed the same amount. In other

words, there is free riding the moment a public good is supplied. Furthermore, Levine takes the argument even further and argues that “Each of us all the time free rides on the culture, knowledge and techniques that are every living human being’s inheritance from preceding generations”.<sup>14</sup> If free-riding is unavoidable in the provision of any public goods and even in any human achievement, we cannot say that there is something wrong with free-riding.

In response to this argument against the free-rider problem, one could accept that it is inevitable that some individuals contribute more while benefiting less and others contribute less while benefiting more. However, it can be argued that some form of contribution is still necessary as compared to no contribution at all. According to Rawls, society is a cooperative venture between free and equal persons for the purpose of mutual advantage. This provides a case for distributing to those who participate in that venture even if they only contribute a little. But an individual who contributes nothing at all cannot be considered part of a community, and in turn cannot receive the benefits of the cooperative venture.

One might then ask: what about the disabled who are incapable of contributing at all? Does the fact that they do not contribute then mean that they do not deserve the advantages of society as a cooperative venture? The disabled are indeed free-riding, but society offers them help in order to meet a human need which is essential

for survival. The need for leisure is not as important as the needs of the disabled. Furthermore, the disabled do not have a choice in whether they can contribute. On the other hand, the idle is aware of the consequences of his actions and should reap the cost and benefits of his actions. Hence, while a disabled person is entitled to receive without contributing, the same cannot be said for someone who chooses not to.

As for Levine's comparison to free-riding on the knowledge of our predecessors, the analogy can be distinguished from the free-riding that occurs when there is a guaranteed income. Free-riding on the knowledge from our predecessors does not inconvenience. But in the case of a guaranteed income, money which could be used to benefit those who contribute is diverted to support those who contribute nothing. Levine's argument that "unless we are prepared to condemn what we cannot avoid, there can be no general proscription of free-riding" does not stand.

The third assumption underlying the free-rider argument is that individuals are entitled to the full produce of one's labor. While individuals deserve to be recognized for their contributions, this is already recognized in the form of the higher income they receive. Individuals owe their earnings to the rest of society as well. For instance, the legal framework set up by society allows for the individual to produce. In the absence of this, "individuals will exert relatively little

effort and a large share of that which is exerted will be devoted to predation and defence".<sup>15</sup> Because the legal framework is owned by the community and a portion of an individual's product is due to the legal framework, the individual owes at least a part of his product to the community.

While this argument does show that the individual is not entitled to the full produce of his labor, it does not show why the produce should go to another person. The individual might owe the produce to the legal framework, but many individuals in society, apart from perhaps judges and policemen, do nothing to contribute to the legal framework. Perhaps this is a case for an individual to pay taxes to protect the legal system, but it does not seem to provide a case for giving money to needy individuals. Even if we accepted that individuals owe a part of their produce to the community as a whole because it "owns" the legal framework, it is unclear whether individuals who do not contribute at all should be regarded as part of the community.

Another argument to show that there is no entitlement to the full product of one's labor could be how what we produce is dependent on the effort of many other individuals in society. Rawls pointed out that a significant part of the social product can be traced to a diversity of talents.<sup>16</sup> However, this argument merely shows that people who do meaningless work might deserve more compensation for their contributions. Guaranteed income is likely to support a group of

voluntarily unemployed people who do not contribute at all to the diversity of talents that Rawls talks about. In fact, guaranteed income will result in those who perform meaningless work supporting those who do not do any work at all.

An argument for redistributing resources to even those who do not contribute would be that every individual has an equal right to the value of natural resources.<sup>17</sup> People ought not to be worse off than they were before the state of nature. In the state of nature, resources were meant for all to share until private appropriation prevented such sharing. As such people with resources owe something to the community, whose members have a natural inheritance to the earth. However, this principle will merely lead to a small handout if we somehow calculated and took the value of natural resources and divided it by the population.<sup>18</sup> This might not be of a value that is sufficient to meet the needs of individuals. To supply a large enough guaranteed income, the remainder will not be from the value of natural resources but the earnings of those who work.

The unfairness of free-riding could be justified if a competing concern outweighed it. For instance, the necessity of providing public goods such as defence or streetlights outweighs the unavoidable free-riding of even those who do not contribute to society. There could even be a case for the need to provide services which are essential for living, such as healthcare. The

question is whether guaranteed income, together with the freedom and leisure that accompanies it, is of sufficient importance to justify free-riding. According to Levine, guaranteed income provides the individual who does not want to work with the means to pursue his conception of the good.<sup>19</sup> The problem with this view is that its logical extension would mean that the government should subsidize any lifestyle which an individual deems as his conception of the good. Levine tries to get around this problem by showing that distinguishing leisure from other expensive goods. Unlike a sea-side condominium, leisure is not substitutable with other consumption goods that can be bought with income. In order to purchase leisure with income, the individual has to work, which goes against his concept of good of being idle.<sup>20</sup> Without a guaranteed income, individuals whose conception of good involves no work are economically coerced into working and thus cannot possibly achieve their conception of good.

One objection is that the pursuit of the conception of the good is by no means an absolute right. It can be curtailed if it went against the interest of society, for instance, if it involved anti-social and illegal behavior. Levine limits the scope to activities that are “compatible with the ideal of a well-ordered liberal constitutional regime”.<sup>21</sup> A life of idleness falls within this scope, but perhaps this scope could reasonably be broadened to include any activity which does not impose an unfair

burden on society to support. In which case, a life of idleness would be considered as a conception of the good which the government does not have to support.

It is unclear when governments need to supplement the means to pursue the conception of good, as opposed to merely not forbidding it by law. An individual might want a harem as a conception of good, but there is nothing wrong with not being able to pursue his conception of good. Many may consider being admired as part of their conception of the good and individuals are permitted to pursue it as long as it is within their means. Like leisure, admiration is a non-substitutable non-commodity good that cannot be bought that could matter fundamentally to an individual. Suppose some individuals have a need for more admiration than other people, it is difficult to say that the government should somehow help provide the means for more admiration. As mentioned above, Levine tries to distinguish leisure from luxuries like a seaside condominium by saying that it goes against a leisure-seeking individual's concept of good if he had to work for income to buy leisure. However, it is not clear why this special property of leisure makes leisure so important as a conception of the good that it justifies imposing a cost on society to support. As a conception of the good, leisure is not substantially different from other luxuries that a government would be unwilling to subsidize.

Allowing individuals as much freedom to pursue whatever their conception of good is might not be a good justification, but a stronger justification would be that guaranteed income corrects an unfair distribution of meaningful activity. Unlike Levine's argument, this is not just about providing more freedom for its sake. Rather it is unfair for some individuals to monopolize meaningful activity at the expense of others. Meaningful activity is fundamental to human life and contributes to the quality of life. Given that work tends to take up a lot of an individual's time, a person with meaningful work would get to spend more time on meaningful activity. Conversely, a person with meaningless work will have little time or even energy for meaningful activity. Because meaningful work is a scarce resource, it is inevitable that some individuals will be deprived of meaningful work even if they all put in as much effort as possible. It is not necessarily true that individuals deserve their meaningful job because they earned it by their efforts. A large part of their achievements are owed to their genes and environment, which they have no control over. It is unfair for some individuals to be deprived of meaningful work because of factors they had no control over. Furthermore, the contributions of those who do meaningful work are dependent on those who do meaningless work. A computer scientist is able to do his more comfortable, fulfilling and better paid work because he does not have to worry about the manufacturing of his clothes or computer parts.

These other tasks are done by factory workers. It is not enough to redistribute work, meaningful activity needs to be redistributed.

A guaranteed income rectifies this unfairness by allowing workers more free time. One option would be for them to leave a meaningless job to devote their time to an unpaid meaningful activity. Another option would be for them to use the increased bargaining power they have from being able to quit in order to shorten their work hours and spend the remaining time on meaningful activity. But what about the employed individual who decide to continue working? The employed individual cannot complain about the unfairness of free-riding because he has the same opportunity to stop work and live off the efforts of those who do work. As he is not forced into working to support the unemployed, it is not unfair that someone else does so. Furthermore, the unfairness of monopolized meaningful activity is greater than the unfairness of free-riding because of the extents they affect life. Free-riding might cause some decrease in the economy's pie, but this is insignificant compared to the amount of meaningful activity that an economically coerced worker is made to give up in order that others can have meaningful work. The improvement to the lives of workers who gain access to more meaningful activity is greater than the loss from decreased contributions.

Apart from the free-rider concern, there is another practical concern: what happens if everyone chooses to

be voluntarily unemployed instead of contributing? In that situation, society will run out of funds to supply a guaranteed income for everybody and the system will fall apart. This is not a likely scenario. As mentioned above, if we accept that human beings need some form of socially productive activity to feel fulfillment, they will continue to work – either in the form of employment or unpaid productive work. Furthermore, it is most likely that people would want to earn more than the basic income in order to consume more.

The more realistic concern is that a guaranteed income results in nobody wanting to do certain jobs in society which are stultifying but nonetheless essential to the functioning of a society. The likely outcome is that the wages for such jobs will rise to the point that people will be willing to take them. This improves the standard of living of workers significantly to compensate for the undesirable nature of their nature. In a situation where we cannot redistribute meaningless activity evenly, we should at least be able to compensate workers who perform the meaningless work substantially for their contributions.

In a more extreme case, it is possible that there is stultifying work that nobody will do the moment there is a guaranteed income because they will only do it if they are economically coerced into doing it. In such cases, coercion might have to be used in order to get people to perform such necessary tasks. While this sounds unpalatable, it needs to be compared to

the status quo where a group of people do such tasks only because they are economically coerced. In either case, coercion will be used. Except under this case, the government has the opportunity to use coercion to ensure that there is equal liability of meaningless work, thus distributing meaningless work in a fairer manner than the status quo which relies on economic coercion. It is unfair for one group of people to spend their entire lives doing meaningless stultifying work. If economic coercion becomes necessary, it might be possible to spread the meaningless work out more evenly across the population such that everybody does a bit of it or takes turns doing it. It must be noted that it is an assumption that the government will distribute meaningless work equally – but at least an egalitarian government would now have the chance to do so. However, in all likelihood, the government will not need to resort to coercion as a huge wage increase should be sufficient to attract people to do even meaningless work.

In conclusion, while free-riding will occur, it is justified as it addresses the unfair distribution of meaningful activity. While guaranteed employment is able to redistribute paid work to help the needy, it does not redistribute meaningful work. It would not be fair for some to be deprived of meaningful activity in order for others to pursue meaningful activity. Guaranteed income provides real freedom to allow an individual to have the opportunity to seek meaningful activity.

### **Works cited:**

- 1 Peter Edelman, “The Big Picture,” in *What’s Wrong with a Free Lunch*, ed. Philippe van Parijs, (Massachusetts: Beacon Press, 2001), p. 104.
- 2 Michael Kalecki, “Political Aspects of Full Employment,” *The Political Quarterly* 14.4 (1943), p. 351.
- 3 Richard Arneson, “Is Work Special? Justice and the Distribution of Employment,” in *Philosophy and the Problems of Work: A Reader*, ed. Kory Schaff (Maryland: Rowman & Littlefield, 2001), p. 202.
- 4 Bernard Cullen, “The Right to Work,” in *Moral Philosophy and Contemporary Problems*, ed. J.D.G Evans (Cambridge: Cambridge University Press, 1988), p.174.
- 5 Ibid. p. 179.
- 6 Richard Arneson, “Is Work Special? Justice and the Distribution of Employment,” in *Philosophy and the Problems of Work: A Reader*, ed. Kory Schaff (Maryland: Rowman & Littlefield, 2001), p. 213.
- 7 Jon Elster, “Is there (or should there be) a right to work?” in *Philosophy and the Problems of Work: A Reader*, ed. Kory Schaff, (Maryland: Rowman & Littlefield, 2001), p. 296.
- 8 Ibid. p. 297.
- 9 Richard Arneson, “Is Work Special? Justice and the Distribution of Employment,” in *Philosophy and the Problems of Work: A Reader*, ed. Kory Schaff (Maryland: Rowman & Littlefield, 2001), p. 213.
- 10 Philippe van Parijs, *Arguing for Basic Income: Ethical Foundations for a Radical Reform*, (London: Verso, 1992), p. 16.
- 11 Andrew Levine, “Fairness to Idleness: Is There a Right not to

Work?," in *Philosophy and the Problems of Work: A Reader*, ed. Kory Schaff (Maryland: Rowman & Littlefield, 2001), p. 325.

12 Ibid. p. 329.

13 Ibid. p. 330.

14 Ibid. p. 330.

15 Philippe van Parijs, *Arguing for Basic Income: Ethical Foundations for a Radical Reform*, (London: Verso, 1992), p. 15.

16 Ibid. p. 16.

17 Ibid. p. 11.

18 Ibid. p. 12.

19 Andrew Levine, "Fairness to Idleness: Is There a Right not to Work?," in *Philosophy and the Problems of Work: A Reader*, ed. Kory Schaff (Maryland: Rowman & Littlefield, 2001), p. 322.

20 Ibid. p. 324.

21 Ibid. p. 321.

## Gettier vs. Justified True Belief

Dominic Sicilian

**Abstract:** This essay is a response to the famous 1963 paper by Edmund Gettier. It offers a thorough interpretation of the "Justified True Belief" model of knowledge, which is used to challenge Gettier's claims, and addresses both of the specific examples given in his paper. It then applies these arguments generally to any conceivable Gettier case and explains why Gettier's objections to the "Justified True Belief" model are not sufficient grounds to abandon it.

## I. Introduction and history

The search for knowledge is one of the most compelling and powerful forces behind the development of humanity. Knowledge is like a crystal of utmost rarity, perhaps the most sought-after commodity in the entire universe, so the quest to acquire it is expectedly daunting. Such a difficult quest is made even more difficult when examined technically, because an obvious observation is made: knowledge is also quite unlike a crystal, in the sense that its definition is not clear, but instead is rather nebulous. This issue of determining what exactly constitutes knowledge was thought to have been conquered by Plato in his *Theaetetus*. Here, Plato introduced the “justified true belief” model of knowledge, which was accepted for over 2000 years. This model defines knowledge as such:

S knows that P if and only if:

1. S believes that P
2. P is true
3. S is justified in believing P

This model requires that for a belief to be considered knowledge, the belief must be true and the believer must have justification for holding the belief. The philosophical consensus was that this view expressed sufficient requirements for knowledge—until 1963 when Edmund Gettier published a short paper on the

subject. Gettier’s paper appeared to reveal problems with the “justified true belief” model.

Gettier’s method of disputing Plato’s definition of knowledge is to offer counter-examples to the view; his examples describe situations in which beliefs appear to fit the “justified true belief” model of knowledge, but do not actually constitute knowledge. His goal in the paper is to prove that Plato’s definition is not sufficient since the cases of non-knowledge appear to fit the definition. Thus, the view that shook the philosophical world was born and dubbed the Gettier problem.

However, there is no need to abandon the “justified true belief” model due to Gettier’s paper. When interpreted fully and correctly, it firmly withstands the Gettier examples.

## II. Justified true belief: a more precise interpretation

“Justification” in the Gettier-interpreted sense is not sufficiently robust to endure Gettier’s counter-examples, which is the core issue with Gettier’s objection; this causes the justified true belief model to appear insufficient as a definition of knowledge. Thus, a complete description of what actually constitutes justification is necessary to explain why the model accurately describes knowledge.

First, a new term to describe a “step” in the knowledge process should be introduced (however, note that it is simply an already-existing step in

achieving “justification” for a belief; this is not an additional condition for knowledge). The new term is “rationale,” and the definition of “rationale” here is the reasoning behind a belief. Every time, in the knowledge-acquiring process, an individual begins by forming a belief. Either simultaneously or in immediate succession, the individual formulates his “rationale” behind the belief. The rationale could be something completely independent of any evidence—for example, an individual could form the belief that eating moth balls is healthy, and his rationale is he likes the design on the moth balls’ packaging; the design on packaging is completely unrelated to whether moth balls are healthy when ingested. The rationale could also be contrary to the evidence—an individual could form the belief that smoking cigarettes is healthy; his rationale could be that it alleviates his stress, and everything that alleviates stress is healthy. When the evidence regarding the health effects of smoking cigarettes is considered, it is obvious that his rationale is not supported. In cases such as these, the rationale does not constitute justification.

To explain why these do not constitute justification, and to explain what does, a second new term should be introduced. This term is “well-founded.” Such a term was used in an account of justification by Richard Feldman, so it is important to note that the use and definition of this term here is different from those pertaining to Feldman’s paper. The definition of “well-

founded” here is to be fully based on — and consistent with — firm, sufficient evidence and truth. The term is applied to an individual’s rationale for his belief. In the previous examples, the individuals’ beliefs had corresponding rationales that were either based on no evidence, or inconsistent with the evidence. Thus, they were not well-founded rationales and therefore did not constitute justification. If, on the other hand, a rationale is well-founded, then it does constitute justification. Justification should be interpreted to inherently require a well-founded rationale. Otherwise, “justification” can be achieved by using insufficient or faulty evidence, and therefore cannot be used to define knowledge.

So far, it can be noted that the only examples of rationales that are not well-founded have involved false beliefs, which means the beliefs never had a chance to constitute knowledge. Now that an account of justification has been established, though, the focus can return to the complete definition of knowledge. The broad “justified true belief” model, as described earlier, can now be elaborated using this new account of justification:

S knows that P if and only if:

1. S believes that P
2. P is true
3. S is justified in believing P, because:

- a. S has a rationale for believing P
- b. The rationale is well-founded

Under this more precisely-defined model, there can be cases in which S believes that P, P is true, and S has a rationale for P—but the rationale is not well-founded and the belief is therefore not justified. For example, S could believe that Earth orbits the Sun; his rationale could be that his cat is gray. Clearly, his rationale is not based on evidence, and thus is not well-founded; this means he has a true belief, but no justification, and therefore does not actually have knowledge in the situation. In another example, S could believe that Earth orbits the Sun; his rationale could be that:

The Sun is one giant proton while the Earth is a giant electron (which is due to his strong apparent evidence for this, which could be because he was told so, both by astrophysicist Neil deGrasse Tyson and by the “science guy” Bill Nye, two of the world’s most reputable scientists), and S was familiar with the fact that an electron orbits a proton in Hydrogen atoms due to his basic childhood/adolescent education in chemistry and physics. Thus, he is led to conclude that Earth orbits the Sun since Earth is an electron, the Sun is a proton, and one electron orbits one proton incredibly often in the universe.

Here, S again has a true belief. This time, he has a more in-depth rationale. However, despite how compelling his rationale may seem (given the

reputability of his sources of information) it is based partially on falsehood (particularly the falsehood that the Sun is a proton and the Earth is an electron); thus, it is not based fully on truth and therefore does not constitute justification. Since the rationale does not constitute justification, the belief (despite being true) does not constitute knowledge. It is simply a case of “lucky” true belief. Just as a belief must be true to become knowledge, a rationale must be well-founded to constitute justification.

### III. Directly refuting Gettier

The examples offered in Gettier’s paper are intended to demonstrate that sometimes “lucky” true beliefs fit the criteria for knowledge under the justified true belief model. However, now that the justified true belief model has been fully described and clarified, the solution to the so-called “Gettier problem” can be revealed. It has been established that “lucky” true beliefs cannot constitute knowledge, due to the strict nature of justification, and the high standards to which a rationale is held. Now, Gettier’s examples can be solved by being shown to not actually constitute knowledge under the justified true belief model.

Gettier offers two cases in his paper. Each case will be examined and refuted. The first case involves a subject Smith, who is competing for a job: Smith believes that the man who will get the job has 10 coins in his pocket. His rationale for believing this is that:

Jones is the man who will get the job (which is due to his strong apparent evidence for this, which could be because he was told so by the president of the company), and Smith was familiar with the fact that Jones had 10 coins in his pocket due to Smith having counted the coins himself. Thus, he is led to conclude that the man who will get the job has 10 coins in his pocket because Jones will get the job and Jones has 10 coins in his pocket.

In the example, Smith ends up getting the job; Smith also unknowingly had 10 coins in his pocket. This means that Smith had a true belief that the man who will get the job has 10 coins in his pocket. Clearly, this is a “lucky” true belief, just as in the final example in section II. Gettier claims that this example involving Smith and Jones fits the definition of knowledge under justified true belief, and since it is clearly not knowledge but instead simply a lucky true belief, the justified true belief model must be flawed. However, according to the fully-elaborated account of the model, as given in section II, Gettier’s case does NOT, in fact, fit the definition of knowledge under justified true belief. This is because Smith’s rationale relies on a falsehood. Just as the final example in section II relied on the false claim that the Sun is a proton and the Earth is an electron, Gettier’s first case relies on the false claim that Jones will get the job. Thus, Smith’s rationale is not well-founded, and therefore does not constitute justification. This means that under the justified true

belief model, Gettier’s first example does not fit the criteria for knowledge and is not a counter-example to the model.

Gettier’s second example is slightly less straightforward than his first. This is due to the use of an “Or” statement. In logic, an “Or” statement is true if one or more of its constituent clauses is true. The case is as follows: Smith believes that Jones owns a Ford OR Brown is in Boston. Here is how he explains this:

Jones owns a Ford (which is due to his strong apparent evidence for this, which could be due to Jones always having owned a car that is a Ford at all times in Smith’s memory, and to Jones having just offered Smith a ride while driving a Ford). Also, if Jones does own a Ford, then any composite “Or” statement that includes “Jones owns a Ford” will be just as certain to be true as the single clause that “Jones owns a Ford,” regardless of how arbitrary the other component(s) are, such as “Brown is in Boston” despite Smith’s complete ignorance of Brown’s whereabouts. Since Smith has strong apparent evidence for believing that Jones owns a Ford, and since “Jones owns a Ford OR Brown is in Boston” is true if Jones owns a Ford, he is led to conclude Jones owns a Ford OR Brown is in Boston.

As it turns out, Jones does not own a Ford, but Brown happens to be in Boston by pure coincidence. Clearly, this is another example of “lucky” true belief. Gettier claims that this example also fits under the definition of knowledge given by the justified true belief model,

and that it therefore a counter-example. Despite the apparent complications introduced in this case that were not present in the previous examples, the solution is still just as straightforward. Smith does have a true belief. However, his rationale is, once again, based on falsehood; just as Case 1 relied on the false claim that Jones will get the job, Case 2 relies on the false claim that Jones owns a Ford. Thus, Smith's rationale is not well-founded, and therefore does not constitute justification. Just as in the last example in section II, as well as in Gettier's Case 1, the individual does not have justification for his belief, and his "lucky" true belief therefore does not constitute knowledge under the justified true belief model. Thus, neither of Gettier's cases are actually counter-examples to the model, which means his paper does not dispute it.

Now that both of Gettier's original cases claiming to be counter-examples have been disputed, it could be asserted that the counter-arguments to Gettier simply exploit particular features that are specific to those examples; after all, there are numerous independently-conceived "Gettier examples" in existence. To claim that Gettier's ideas and points have been completely discredited due simply to the rejection of only two of the countless "Gettier examples" in existence would be overambitious and questionable. These rebuttals from a defender of Gettier would be valiant and seemingly legitimate, and perhaps could give the appearance of a stalemate in the argument.

Responding to an opponent making such claims would be quite simple, however. Every "Gettier example" relies on the same feature as Gettier's cases; a "lucky" true belief with a rationale based on falsehood or otherwise insufficient evidence. In all such cases, the rationales, as shown in sections II and III, do not constitute justification, and the beliefs therefore do not constitute knowledge under the definition given by the justified true belief model. Furthermore, even more generally than that, not only can it be contended that no true counter-examples have been raised thus far to dispute the justified true belief account, it can also be contended that no counter-examples can ever exist that successfully dispute the account in its form as described by section II. There cannot be a belief that is true, whose rationale is based only and completely on truth and sufficient evidence (which makes it well-founded and therefore constitutes justification), whose apparent knowledge is revealed to be simply "lucky" true belief. Given the factual nature required of the rationale—the truth of which being as important to having knowledge as the truth of the belief itself—to constitute justification, it is not possible for a true belief that is justified (i.e. a true belief, supported by truth) to be simply "luck." In cases where there is both a true belief, and a rationale constituting justification, the belief will always legitimately be knowledge, and will never be the product of "lucky" true belief. These

responses to the Gettier supporter nearly finish the solution to the Gettier problem.

The final objection to these arguments against Gettier would be to disagree with the definition of justification itself. In such an argument, the opponent would say that falsehoods can justify beliefs, and would point, yet again, to the Gettier examples. The opponent would insist that the beliefs in those examples are justified. Perhaps he would also propose an example of his own: Q believes that X is Sicilian, and the belief is true. Q's rationale is:

Everyone Q has ever seen that looks like X has been Sicilian (which is based on each person's testimony that he/she is Sicilian). Q is led to conclude that X is Sicilian, too.

However, the people on whose testimony the rationale was based were intentionally fooling Q; they are all actually from Calabria, an Italian region north of Sicily, which means they are not Sicilian. The opponent, after giving the example, would then say that Q is justified despite his rationale being based on falsehood. The quick response to this opponent's example is to point out that it is simply another Gettier case, and has therefore already been thoroughly refuted. However, to complete the argument, a more general explanation must be given to this opponent to explain why Q is not justified.

In the Gettier examples, or in the "Sicilian" example at hand, it may seem intuitive to believe that the

subjects are justified due to their strong apparent evidence, regardless of whether their rationales are based on truth; the rationale seems strongly supported and reasonable enough to constitute justification. This general point, and not the specific "Sicilian" example, is the true argument of the final opponent; this is perhaps the strongest argument against the argument for "justified true belief." However, it can be dispelled too. Consider an example in which Y believes that apples are healthy; assume this is a true belief. Y's rationale is that apples are manufactured by wizards using magical unicorn dust that is 100% beneficial to the human body when ingested. Clearly, Y is by no means justified, since his rationale is based on blatant falsehoods; even the opponent at hand would agree, and likely would say that such an absurd example does not apply to his argument. In fact, it does. The reason Q is not justified in the "Sicilian" example, and more generally the reason justification cannot be based on falsehood, is as follows. If it is obvious and easily agreed-upon that Y is not justified in the "apple" example due to his rationale being based on falsehoods, then it must be accepted that falsehood-based rationales can never constitute justification. If one does not accept this, then his account of knowledge will be compromised, because subjectivity will thus become a feature of it; subjectivity will play a major role in deciding when falsehood-based rationales are acceptable and when they are not. The acquisition of knowledge is inherently an objective

enterprise; therefore any account of knowledge that includes subjectivity as a feature is inherently compromised. To clarify this point further still: the definition of knowledge must be universally applicable to correctly and consistently identify knowledge, and if the application of a given account of knowledge varies between subjects, then that account is therefore not universally applicable, which makes it faulty, and thus it cannot be accepted as the definition of knowledge. Therefore, justification cannot be based on falsehood, and the response to this final opponent is complete. These final arguments thus solidify the objections to Gettier, and therefore complete the solution to the infamous Gettier problem.

#### IV. Conclusion

The justified true belief model was the accepted account of knowledge for over 2000 years. The publication of Gettier's paper appeared to offer counterexamples and thus successful objections to the model, which led to the philosophical consensus to be in favor of Gettier and in opposition to the justified true belief model. However, after more than 50 years of the new consensus, it has now been contended and shown that the justified true belief model is a legitimate, sufficient, and successful account of knowledge, just as it was once thought to be.

## Am I Doing Bad Science?

Margaret Min

**Abstract:** The legitimacy of a scientific theory seems to depend heavily on whether or not that theory has made accurate predictions. For example, consider the theory that claims the misuse of antibiotics, like penicillin, can lead to potentially deadly antibiotic resistance. This theory was based on observations of the development of antibiotic resistance in bacteria, yet there was resistance to this theory until the theory proved itself many times over with confirmed predictions (i.e. many predicted instances of strains of bacteria evolving resistance to antibiotics). Today, the theory that misuse of antibiotics can lead to antibiotic resistance is widely accepted in the scientific community. More generally, most scientific theories are judged at least in part on whether they have predicted a large amount of data, to the point where it is necessary for theories to have predicted data to be taken seriously at all. Interestingly, pre-existing data that is explained by a new theory is typically viewed as less supportive of that theory than data that has been predicted. This paper will critique the widespread belief that predicted data is better than pre-existing explained data by reviewing two philosophers' opinions on this matter, Peter Lipton and Carl Hempel.

## I. What are we arguing about?

This paper will address two of the arguments Peter Lipton presents in his paper, “Testing Hypotheses: Prediction and Prejudice”. Before presenting those arguments, let me first explain the background information related to Lipton’s arguments. In postulating scientific hypotheses, there are two main types of data used. In this paper I will call them accommodated data and predicted data.

Accommodated data are data that are collected before a hypothesis is proposed. Typically, scientists will construct scientific hypotheses based on their interpretations of accommodated data. For example, a botanist who is battling an infestation of ants in his greenhouse using bug-spray might notice that the plants on which he sprays bug-spray tend to produce fewer seeds. The botanist observes 100 of his plants, 67 of which were sprayed with bug-spray and 33 of which were left alone. All 67 of the bug-sprayed plants produced fewer than half the number of seeds produced by the 33 non-bug-sprayed plants. From this (accommodated) data, the botanist proposes his scientific hypothesis that bug-spray has toxic effects on plants and makes them produce fewer seeds than normal. (Importantly, accommodated data is labeled as “accommodated” only for the role it plays in postulating scientific hypotheses. The hypothesis attempts to explain and account for the data, in other words, accommodate the data. However, nothing about

accommodated data taken alone should suggest that it is being accommodated or accommodating something else.)

Predicted data are data that are collected after a hypothesis is proposed. After the botanist proposes his hypothesis, he predicts that if he were to plant the seeds of the bug-sprayed plants in a no-bug-spray-zone, effectively eliminating the bug-spray factor, normal seed production would be restored in the next generation. He carries out this experiment and observes that his prediction was correct. The predicted data are his observations of the results of his experiment.

Generalizing to all scientific hypotheses, we see that all scientific hypotheses are supported by accommodated data (that is how hypotheses get proposed in the first place), while a majority of those scientific hypotheses are additionally supported by predicted data. A small subset of scientific hypotheses is supported only by accommodated data and not by predicted data (presumably, these would be very young hypotheses that have yet to be tested by experiments).

It is a widely accepted view that predicted data provide stronger support for hypotheses than do accommodated data. Lipton agrees with this intuition and argues for two versions of this view, what he calls the weak advantage thesis and the strong advantage thesis. In this paper, I will present Lipton’s weak advantage thesis (II) and explain my agreement with

it (III), followed by a presentation of Lipton's strong advantage thesis (IV) and my disagreement with it (V).

## II. Lipton's weak advantage thesis

One of the virtues of scientific hypotheses is having precise and relevant supporting data. This is an evidential virtue, meaning that it is a virtue of a hypothesis in nature of how well it is supported by empirical evidence. Peter Lipton, along with many scientists I take it, would agree that having precise and relevant supporting data is very important for scientific theories. A hypothesis that has weak supporting data is a worse scientific theory than a hypothesis that has strong supporting data, all other things being equal. Although Lipton does not go into too much detail in his paper about why he believes this is an evidential virtue, I would guess that he believes it in a simple and intuitive way as follows. Consider the botanist's hypothesis in Part 1 that states that bug-spray has toxic effects on plants and makes them produce fewer seeds. And let's say that instead of his actual data (observations of plants sprayed by bug-spray or not sprayed), the botanist instead collects his data from observations of goats that were sprayed with bug-spray. He observes that goats sprayed with bug-spray produce fewer offspring than goats that were left alone. Though this is an interesting phenomenon, observations of goats certainly provide less support for his hypothesis about plant seed production than would have observations

of actual plants. And all other things being equal, for any given hypothesis, we should prefer more strongly supportive data to weakly supportive data because this preference will yield better scientific theories.

Considering this fact, that a hypothesis is more strongly supported by precise and relevant data, Lipton presents his weak advantage thesis as follows:

Weak advantage thesis: Predictions tend to provide stronger support for hypotheses than do accommodations.

Lipton believes the weak advantage thesis because successful predictions tend to be more precise and relevant data than do accommodated data. Lipton argues that in typical scientific practice, a shrewd scientist will make observations on general phenomena and pick out just one or a few statistically significant patterns (i.e. a pattern that does not occur merely by chance) from the vast data. From these patterns, the scientists will make a hypothesis. With the hypothesis, the scientists will create experiments to further test the hypothesis. Importantly, these experiments are designed specifically to test the hypothesis in question, so the data collected (predicted data) will directly prove or disprove the hypothesis. In contrast, the original observed phenomena used to postulate the hypothesis (accommodated data) were not designed to answer any specific questions and much of the observed data may in fact be completely irrelevant to the resulting hypothesis except for select data points. Thus, Lipton

argues, predicted data tend to provide more powerful support for hypotheses than accommodated data because the way in which that data is collected. The passive way of collecting accommodated data can result in too much noisy and irrelevant data while the aggressive and specific experimentation done to collect predicted data will yield more useful and telling results.

Importantly, Lipton recognizes that these patterns of predictions providing stronger support than accommodations are not going to show up in every scientific theory. There may be cases in which the accommodated data is more supportive of the theory than predicted data. Perhaps the accommodated data came from observations of some phenomenon that cannot be repeated in artificial experiments. All the scientist can do is create models of the original phenomena. In this case, it seems less clear that the predicted data collected from these model experiments provides stronger support than the accommodated data collected from observations of the original phenomenon. These types of situations can occur, but Lipton claims only that predictions tend to provide stronger support than accommodations, and crucially not that predictions always provide stronger support than accommodations.

### **III. Agreement with the weak advantage thesis**

I agree with Lipton's weak advantage thesis. To illustrate, let's return to the botanist from Part I. Let's

tweak the situation a little and say that the botanist attempted to control his ant problem by spraying bug-spray on the plants, lowering the temperature of the greenhouse (ants prefer warm temperatures), and caressing the leaves of the plants to encourage them to stay strong in the face of the ant armies. The botanist records which plants he sprays bug-spray on, but does not record the plants he caresses or the temperature change in the greenhouse. Still, he makes his observations of 100 plants and sees that the 67 plants he sprayed bug-spray on produced fewer seeds than the non bug-spray plants. Using this data, he postulates the same hypothesis and designs further experiments to test it. Realizing that there are other potential variables (greenhouse temperature and relative botanist love/caressing), he designs further experiments to control for these variables to test only the effect of bug-spray. And what do you know his predictions are still correct! Bug-spray is the culprit of lower seed production, not lower temperatures or his excessive plant love. With this example, we can see that the accommodated data was certainly enough for the botanist to propose his hypothesis, but it was considerably weaker in support than the predicted data. The accommodated data was weaker simply because it was too unfocused, there were too many other factors that could have influenced the data. The predicted data was collected in a way that was designed specifically to eliminate those factors so

that the hypothesis at hand could be answered more definitively.

#### **IV. Lipton's strong advantage thesis**

Lipton's strong advantage thesis appeals to a different sort of virtue of scientific hypotheses, the theoretical virtue of simplicity. Ockham's razor is a general principle of theory selection that says that simple theories should be chosen over more complex theories (all other things being equal) because simpler theories are more likely to be true. This paper will not delve too deeply into the question of why simplicity plays a factor in choosing scientific theories (that would constitute its own paper!) and will instead present a simple example to show why Ockham's razor seems to be true in at least one way. Consider two explanations for why humans have 10 fingers. One explanation is that fingers evolved through 200 genetic mutations occurring over the last 100 million years. Another explanation is that fingers evolved through 2,000 genetic mutations occurring over the last 100 million years. Ockham's razor says that we should choose the first hypothesis because it is simpler - it requires 1,800 fewer genetic mutations to reach the same outcome. Considering the fact that genetic mutations are very rare, and genetic mutations specifically towards finger evolution are even rarer, the first hypothesis, statistically speaking, is more likely. The first hypothesis requires fewer really rare occurrences to occur. More generally, theories

that minimize the need for really rare occurrences to occur are simpler and more likely. This is only one of many types of simplicity. Some types of simplicity can conflict with each other, others are difficult to judge, and still more are hard to weigh against other types of simplicity.

Another virtue of scientific theories that plays a part in Lipton's second argument for the strong advantage thesis is the evidential virtue of having wide and diverse data supporting the hypothesis. To illustrate, consider a scientific hypothesis stating that the more chickens my grandmother raises, the more frequently she will eat omelets for breakfast. And let's say that the scientist that postulates this hypothesis has 2 data points to support it. When my grandmother has 15 chickens, she eats omelets 5/7 days of the week. When my grandmother has 16 chickens, she eats omelets 6/7 days of the week. The scientist reasons that the data points show that when we add one more chicken to my grandmother's family, she eats one more omelet every week and thus, the more chickens my grandmother has, the more omelets she will eat. However, the scientist's hypothesis is poorly supported because two data points is not enough to prove that more chickens means more omelets. For one, statistically speaking, we have no idea if these data points actually show a trend, or are just a result of chance fluctuations. Secondly, there may be other patterns of chickens and omelets that cannot be revealed by only 2 data points. For instance,

it could be that if my grandmother has fewer than 10 chickens, she will not eat any omelets, if she has 10-20 chickens, she'll eat 5 or 6 omelets a week, but any more than 20 chickens and she'll only eat fried chicken (i.e. no omelets). A trend like that can only be revealed by testing my grandmother extensively in each of those ranges and will not be discernible from a data set with only 2 points. Generally, having wide and diverse data will provide greater support for a hypothesis because it will more likely reveal consistent trends and thus lead to a hypothesis that is more accurate.

Lipton claims that there is a conflict between the theoretical virtue of simplicity and the evidential virtue of having wide and diverse data. In science, the ideal is to have both wide and diverse data, but also to have a simple theory. This can be challenging for scientists because the wider and more diverse their data is, the harder it is to keep the theory simple while still accounting for all that data. Thus, scientists are often required to make compromises between the two virtues, partly sacrificing a very simple theory as well as not being able to account for all the data at hand to meet at some happy middle. Lipton argues however, that scientists, being imperfect humans, are not very good at judging how much simplicity to sacrifice and how much data to sacrifice. Most scientists tend to try to keep all the data they can ("over-fitting" their data), sacrificing the simplicity of their hypotheses. Lipton calls this "fudging a hypothesis". Thus, many scientific

hypotheses are more complex than they need to be, and by Ockham's razor, they are thus less likely to be true than a simpler hypothesis. This brings us to Lipton's 2nd argument, what I will call his fudging argument, which supports the strong advantage thesis as follows.

Strong advantage thesis: A single, particular observation that was accommodated would have provided more support for the hypothesis in question if it had been predicted instead.

Why does Lipton think the strong advantage thesis is true? Lipton thinks that it is important to avoid making the common mistake of fudging a hypothesis (i.e. trying to explain too much of the data), for if we are able to avoid this mistake, we will have a simpler and thus more accurate theory. Predicted data arrive on the scene after a hypothesis or theory is already postulated. Thus, a scientist cannot fudge her hypothesis to fit more of her predicted data because the predicted data didn't exist when she was making the hypothesis. She wouldn't even know how to fudge her hypothesis to over-fit the predicted data. The scientist will only make the mistake when it comes to accommodated data, for that is the data used to create hypotheses. Thus, scientists will always run the risk of fudging their hypotheses since all scientific hypotheses are based in part on accommodated data. Scientists that have more accommodated data will be at a higher risk of fudging their hypotheses because there is more data available that they might want to explain by fudging hypotheses.

The strong advantage thesis is true because if the same hypothesis has less accommodated data than the other, but on total the same amount of supportive data (accommodated data + predicted data), then it is more likely to be a more accurate hypothesis because the scientist postulating that hypothesis had a lower risk of fudging it.

## V. Disagreement with the strong advantage thesis

### i) Disagreement #1

Lipton's support of the strong advantage thesis stems from his view that science should make a compromise between postulating a simple theory and accommodating for the most data they can, but human scientists are prone to imperfection at achieving this balance. Though I do agree that simplicity plays an important part in science, it seems that science should first prioritize explaining data. Science is in the business of explaining what the world is like and why the world is like that. And as humans, our only access to the world is through our experiences. In science, these experiences translate into our data and become our only resource in making any claims about the world. If scientists want to do good science, they should account for their data.

However, Lipton argues that we should actually sacrifice some of our data in the interests of simplicity. He says that data can often times be too wide and varied

to be accounted for in a simple theory. I would argue that though simplicity should still be kept in mind (for a very complex theory would be useless to us), the first priority in science is to explain the data. Simplicity should only come into play when comparing two hypotheses that explain the data equally well. Lipton calls it fudging hypotheses to explain more data, but his process of fudging hypotheses actually seems like good scientific practice. If a scientist has to trash some of her data because it doesn't agree with her simple theory, then she just hasn't found the best explanation for the data.

Crucially, I do not support the over-fitting of data. Instead, I value good statistical analyses that account for chance fluctuations in the data. When scientists appear to over-fit their data, it is not because they are trying to explain too much of the data in their theory. The problem instead lies in the fact that they have not used statistical analyses to explain data points that seem to deviate from trends. A scientist that employs effective statistical analyses will be able to explain all of her data while avoiding over-fitting.

Lipton could argue in response that even the use of statistical analyses is subject to the imperfection of humans, so a scientist that does statistical analyses will still run the risk of over-fitting her data if she does her statistical analyses inadequately. And because she runs this risk, it's ultimately better to just have more predicted data and less accommodated data since the

former will not pose this risk of over-fitting. I would argue in response that a scientist who does inadequate statistical analyses would find both accommodated data and predicted data to be problematic for a theory. On the one hand, doing inadequate statistical analyses on accommodated data may result in a hypothesis that over-fits the data, and that's obviously bad. But additionally, if the scientist is doing inadequate statistical analyses on predicted data, then she will likely make flawed interpretations of her data. She might take a certain data point that deviates from her hypothesis as evidence to reject her hypothesis, even if that data point was just a chance fluctuation in the experiment that could have been accounted for in statistically analyses. Or, she might look at her scattered data points and conclude that the empirical evidence supports her hypothesis, even if statistical analyses would show that the data actually indicates a different pattern. Thus, the fact that humans can be imperfect in carrying out statistical analyses poses a problem for scientific practice on the whole including both accommodated data and predicted data. The problem isn't when the data is collected, but that the scientist just doesn't know how to interpret the data. And so this human imperfection should not be a reason for us to prefer predicted data to accommodated data.

ii) Disagreement #2

It seems that even if Lipton's reasoning of fudging hypotheses were good reasoning, it would still fail

to support the strong advantage thesis. The strong advantage thesis focuses on how well a particular piece of data supports a hypothesis, that support depending on whether the data was found before or after the hypothesis is postulated. Lipton argues that the strong advantage thesis is true because hypotheses supported by fewer accommodated data are less likely to be fudged. His reasoning focuses on the chances an imperfect human scientist will fudge his hypothesis, and crucially does not focus on how supportive a piece of data is if it is accommodated or predicted. The contrast between accommodated and predicted data is revealed only coincidentally by how the two types of data happen to influence the likelihood of fudging. If we could build a machine whose only function is to balance simplicity and explaining the provided data to the exact levels Lipton would want (eliminating human imperfection), it seems like his fudging argument would no longer apply. However, the strong advantage thesis would still apply because hypotheses generated by this machine would be supported by varying amounts of accommodated data and predicted data. If Lipton wants to support the strong advantage thesis, then he will need to come up with an argument that directly addresses the difference in supportiveness of accommodated data and predicted data.

iii) Disagreement #3

Lipton's support of the strong advantage thesis relies on the imperfection of humans in being able

to construct simple theories that also accommodate for most of the data. I would like to focus on the imperfection of humans. My disagreement will first appeal to an argument made by Carl Hempel in his book *Philosophy of Natural Science*. Hempel says, "...the strength of the support that a hypothesis receives from a given body of data should depend only on what the hypothesis asserts and what the data are." In contrast, Lipton seems to argue in his fudging argument that the strength of support a given piece of data provides for its hypothesis depends on the likelihood that the imperfect human scientist has fudged his hypothesis to fit that piece of data. And Hempel (and I) would argue that that seems wrong. Science should be objective and certainly, the relation between theories and data should be objective. For any given hypothesis, how well it is supported by data should not depend on who proposed the hypothesis and how good they are at doing good science. Let's say Mary is really great at doing science (in at least the way Lipton thinks), but Jane has really bad judgment and often compromises simplicity to over-fit her data. Given a hypothesis A, supported by data points 1-10, how strongly those data points support hypothesis A should not depend on whether Jane or Mary is the scientist collecting the data and postulating the hypothesis. The "who" seems irrelevant to science. The fudging argument does not seem to be consistent with the objective nature of science.

Hempel (and I) would object to Lipton's fudging argument and his support of the strong advantage thesis in another way as follows. The differential valuing of predicted and accommodated data is misguided. Both types of data can provide equal support to any given hypothesis. How much support a piece of data can provide depends only on what the hypothesis asserts and what the data is, and crucially, does not depend on when the piece of data is collected. Hempel says that one way to illustrate this is to consider the spectral emission lines produced by certain gases, a phenomenon called the Balmer series. These emission patterns are described by a mathematically elegant and simple formula constructed by J. J. Balmer. Hempel would argue that when it comes to mathematically elegant hypotheses like this one, Lipton's fudging argument does not apply for there is no conflict involving simplicity, as the hypothesis is already very simple. Mathematically elegant hypotheses like this one are not over-fitting their data and are not fudged, even if those hypotheses happen to be constructed from lots of accommodated data. Thus, by Hempel's and (my) view, Lipton should modify his fudging argument in favor of the strong advantage thesis such that it only applies to certain types of hypotheses, that is, hypotheses that are difficult to assess in terms of complexity, hypotheses that are not or cannot be simplified into mathematically elegant formulas. It is not the time at which data is collected that gives rise to

different levels of support of a hypothesis, but rather the content and form of the hypothesis and the data.

## VI. Concluding remarks

My paper has hopefully illuminated some of the problems that arise in trying to justify the widely accepted notion that data that has been predicted provides stronger support for a hypothesis than data that has been accommodated, all else being equal. Many philosophers and scientists, like Lipton, point at the seemingly obvious risk of over-fitting accommodated data by fudging hypotheses to explain more of the data. However, explaining all of the data should be a goal of science, not something to be avoided. There is a risk of over-fitting accommodated data, but there is a similar risk of wrongly confirming or rejecting hypotheses using predicted data. And so it isn't clear why we should prefer predicted data to accommodated data. Finally, it seems like scientific practice should be objective. When scientists make hypotheses from their data, or use data to confirm or reject their hypotheses, it doesn't seem like the support of a hypothesis depends at all on whether the scientist is more or less likely to do bad science. So any argument that claims that accommodated data is worse than predicted data because accommodated data make scientists more likely to do bad science will not suffice because it will not be consistent with the objective nature of science.

## Pornography as Group Libel

Irmak Aydemir

**Abstract:** This paper argues that pornography is an instance of group libel against women. Using Jeremy Waldron's framework of hate speech, the paper defines group libel as undermining the reputation of a group and the social standing of its members, and identifies four different types of group libel. The paper argues that pornography is group libel because it disseminates misogynist ideals: first, by presenting women as essentially sexual; second, by commercializing and illegitimizing their emotions; and third, by depicting women as naturally and willingly sexually subordinate to men. Each of these arguments are linked to the four types of libel and it is asserted in closing that, as an effect of pornography, the status of women as equal citizens in society is diminished.

Jeremy Waldron states two reasons to regulate hate speech. First, he says, inclusiveness is a public good, and our society “sponsors and ... is committed to” this public good.<sup>1</sup> Groups in diverse multicultural landscapes, Waldron argues, “must accept that society is not just for them; but it is for them too, along with all of the others”.<sup>2</sup> Members of groups do not deserve to “face hostility, violence, discrimination, or exclusion by others” on the basis of their group membership, yet where hate speech is tolerated, inclusiveness is undermined. Hate speech is an “environmental threat to social speech, a sort of slow-acting poison, accumulating here and there, word by word”.<sup>3</sup> The threat is to the public good of inclusiveness, but it is also to the individual dignities of persons in society who are targeted by hate speech. A second reason to regulate hate speech, then, is to preserve the basic reputations of individuals in society; it is to allow each member of society to rely implicitly on the idea that they may go about their daily lives in peace. Waldron’s project focuses mainly on ethnic and religious injustices: anti-black and anti-Muslim propaganda tend to illustrate his points. Nonetheless, Waldron’s “modest intention”—to make a case for “the constitutional acceptability of [hate speech] laws”—is equally appropriate for matters of sexism.<sup>4</sup>

In this paper I use Waldron’s framework of hate speech to argue that pornography is an instance of group libel. I proceed, first, by explaining Waldron’s

choice of terminology to make clear what is meant by group libel. Second, I list the four ways group libel takes form to determine the focus of my investigation. I also briefly speculate on the reasons why anti-pornography legislation tends not to gain supporters in order to preemptively dispel some objections to my task. Third, I argue that pornography undermines the dignity of women by disseminating dangerous misogynist ideals. The argument proceeds by clarifying the term “misogynist ideal”, arguing for the female gender itself as such an ideal, and identifying misogynist ideals universally present in pornography. Fourth, I expound the implications of the view that pornography disseminates misogynist ideals to prove that pornography takes the forms of group libel outlined by Waldron. In conclusion, I assert that pornography should be regulated as group libel.

“Hate speech” is a problematic term for two reasons. First, the word “hate” implies that we are interested in correcting negative emotions, silencing certain viewpoints, or characterizing hate speech issues as attitudinal issues. Second, the word “speech” suggests that we are looking to regulate the spoken word or sets of politically incorrect vocabularies. What we are really after with hate speech legislation, thinks Waldron, are restrictions on tangible forms of hateful message. “Hate” does not refer to the speaker’s motivations, but rather to the possible effects of his speech. “Speech”

does not refer to the speaker's conversation, but rather to the speaker's expression of a viewpoint.

Waldron is particularly concerned with messages that become "established as ... visible or tangible feature[s] of the environment—part of what people can see and touch in real space (or in virtual space) as they look around them".<sup>5</sup> He believes hateful messages of this kind pose a special threat to public order by compromising the aesthetic of a well-ordered society. A society that allows hate speech publications, argues Waldron, hoards or festoons its lampposts with "depictions of members of racial minorities [characterized] ... as bestial or subhuman". It celebrates and excuses "genocidal campaigns of the past" with images of the swastika. It permits posters baselessly defaming members of minority groups as "criminals, perverts, or terrorists".<sup>6</sup> It allows, in short, the semi-permanent and public expression of hatred. It allows the sending of messages that become "part of the permanent visible fabric of society so that, for the [Muslim] father walking with his children ... there will be no knowing when they will be confronted by [anti-Muslim] signs, and the children will ask him, 'Papa, what does it mean?'"<sup>7</sup>

The term "group libel" captures the phenomenon Waldron has in mind. "Group libel" and "group defamation" are interchangeable. The terms refer to "attacks on human dignity by insulting, maliciously maligning, or defaming part of the population".<sup>8</sup> For

Waldron, criminal law is concerned with group libel insofar as it is concerned with vindicating public order. Laws are concerned "not just [with] pre-empting violence," he says, "but [also with] upholding against attack a shared sense of the basic elements of each person's status, dignity, and reputation as a citizen or member of society in good standing".<sup>9</sup> I argue for pornography as group libel instead of hate speech to capture this fact: that citizens of societies have a shared sense of what it means to be a member of society in good standing. Nonetheless, where I quote authors who use the term "hate speech" instead, I understand the terms to mean the same thing.

What does group libel do, and how? Group libel undermines the reputation of a group and thereby the social standing of individuals belonging to that group. Waldron says the concern is not about the group as such, "as one might be concerned about a community, a nation, or a culture".<sup>10</sup> The concern, he says, is ultimately individualistic. Group libel laws "look ... to the basics of social standing and to the association that is made ... between the denigration of that basic standing and some characteristic associated more or less ascriptively with all members of the group".<sup>11</sup> These associations are made "in the hate speech, in the libel, in the defamatory pamphlet, poster, or blog".<sup>12</sup> They are the ideas formed about individuals in virtue of their group membership.

Group libel takes four main forms. The first form I will call libel by factual claim. Waldron says this form of libel proposes factual imputations about a group in order to have these imputations “accepted at a general level”.<sup>13</sup> An example is the 1952 case of *Beauharnais v. Illinois*. Joseph Beauharnais, the founder and director of the White Circle League of America, was charged for claiming African-American people “mongrelized” the white race with “rapes, robberies, guns, knives, and marijuana”.<sup>14</sup> The courts held that Beauharnais had portrayed an entire class of citizens as depraved, criminal, unchaste, and unvirtuous. The effects of such claims, argues Waldron, have profound effects on all members of the group. A person who belongs to a group vilified in this way is denied what dignity and respect she deserves: her identity as a member of the group has a direct impact on her social standing as a member of society.

The second form of group libel I will call libel by denigrating claim. This libel characterizes a group on the basis of false or derogatory opinion. It aims to damage the social and cultural reputation of a group so that individual members of the group are isolated and stigmatized. The case of *R. v. Keegstra* was libel by denigrating claim. James Keegstra, a high school teacher, taught his classes that Jewish people sought the destruction of Christianity and invented the Holocaust to gain sympathy. Catherine MacKinnon characterized this type of libel as promoting the disadvantage of a

historically disadvantaged group to “shape their social image and reputation, which controls their access to opportunities more powerfully than their individual abilities ever do”.<sup>15</sup>

Libel by dehumanizing claim is what I will call the third form of group libel. This libel “[damns] the members of the group with vicious characterizations that dehumanize their ascriptive characteristics and depict them as insects or animals”.<sup>16</sup> Waldron uses the example of the “racist agitator” who depicted Britons of African descent as apes in the late 1970s. This type of libel “denigrate[s] [one’s] fellow citizens in bestial terms”. It destabilizes the belief that “all humans, whatever their color or appearance, are equally persons, with the rights and dignity of humanity”.<sup>17</sup>

The final form of group libel I will call libel by discriminatory claim. This form of libel denigrates the members of a group “by embodying slogans or instructions intended implicitly to degrade (or signal the degradation of) those to whom they are addressed”.<sup>18</sup> Any sign, emblem, or visual publication that reads “Muslim Out!” or “No Blacks Allowed” is libel by discriminatory claim. This libel displays a “poisoning of the social environment” that declares racial or religious hostility toward a group.<sup>19</sup>

These are the four ways group libel can be identified. Waldron believes these attacks are attacks on a person’s dignity. He understands dignity as “basic social standing”, as “the basis of [one’s] recognition as

[socially] equal and as [a bearer] of human rights and constitutional entitlements”.<sup>20</sup> Dignity is a social and legal status that must be upheld by society and the law.<sup>21</sup> It differs from “subjective aspects of feeling, including hurt, shock, and anger” because dignity “has to do with how things are with respect to [someone] in society, not with how things feel to them”.<sup>22</sup> This is an important point: I am concerned not with how pornography makes women feel, but rather with how it affects their status as equal citizens in society. Groups have the right to feel assured that “all are equally human, and have the dignity of humanity, that all have an elementary entitlement to justice, and that all deserve protection from the most egregious forms of violence, exclusion, indignity, and subordination”.<sup>23</sup> “Assurance” here is understood as a technical term, not one referring to one’s emotional state of mind, but rather to one’s social standing as a citizen. This assurance is “implicit, as though the underlying status of each person as a citizen in good standing goes without saying”.<sup>24</sup>

My task is therefore to prove that pornography is an attack on the dignity of women, an attack that undermines their right to this assurance. My challenge is to reveal the ways in which pornography takes the various forms of group libel outlined above. The examples I provided for each form of group libel were clear. They were obvious cases of libel because their messages, effects, and relationships to the respective libellous forms were candid and largely undebatable.

Pornography is a trickier case. Its messages and effects are the subject of debate even among feminists.

I am interested in these academic debates, but partial to everyday debates that inform our pre-theoretical opinions about pornography. This is because pornography exists in real space, and everyday citizens feel and live its effects. With this consideration in mind, I believe there are three main reasons why people hesitate to support anti-pornography legislation.

One reason is that heterosexuality, with its heteronormative mandates, is a dominant but concealed institution, one that Adrienne Rich calls a “beachhead of male dominance”.<sup>25</sup> Heterosexuality as an institution covertly permeates all aspects of female life. It convinces women that marriage and sexual orientation towards men are inevitable; it bases female economic survival on traditional roles of femininity; it subordinates women under men. If Rich is right, the effect this institution has on the psyche and identity of women is fundamental. Pornography, as equipment for male dominance, can be understood as simply reiterating society’s foundational social dynamics. It is hard to criminalize the negative principles of pornography if those principles structure our everyday lives.

Another reason is that, historically, sexist claims have been understood as claims to objectivity. The idea that women are naturally mothers, for instance, has mostly gone unchallenged. Whatever sexism

pornography is charged with is dismissed as an attempt to pervert the facts. “Men do not objectify women,” someone argues. “Men just are like that. They enjoy pornography because they are naturally sexual and visual creatures.” These “facts” are dubious. As Joan W. Scott points out, “part of the project of some feminist history has been to unmask all claims to objectivity as an ideological cover for masculine bias by pointing out the shortcomings, incompleteness, and exclusive of ‘mainstream’ history”.<sup>26</sup> The idea that there exist naturally-occurring distinguishing facts about the genders—and that this has “always been the case”—is itself a sexist interpretation of history. The negative effects of pornography are impenetrable from this perspective.

The final reason is that women are, and have historically been, socialized to internalize sexism. Social mechanisms regulate, and have historically regulated, our cultural ideas of men and women’s personalities, bodies, and roles. Susan Bordo discusses one effect of this social regulation. She argues that “the body—what we eat, how we dress, the daily rituals through which we attend to the body—is a medium of culture”.<sup>27</sup> It is the site where “prevailing forms of selfhood, desire, masculinity, [and] femininity” are impressed.<sup>28</sup> It is no accident, she argues, that women are more likely than men to develop eating disorders: “in the pursuit of slenderness and the denial of appetite the traditional construction of femininity intersects

with the [social] requirement for women to embody the ‘masculine’ values of the public arena”.<sup>29</sup> It is taken for granted that society’s beauty ideals are somehow objective or obvious. Men and women often tacitly strive toward the ideals without self-reflection. This lack of self-transparency characterizes internalized sexism. Anti-pornography legislation loses support where individuals fail to reflect on the origins of their views on sex, sexuality, and gender.

These three considerations were intended to anticipate and undermine common reasoning that underlies objections to anti-pornography legislation. Of course, there are dozens of other possible objections. I chose to address objections of this type in order to make explicit the assumptions underlying my case for pornography as group libel. I am assuming, first, that there exist patriarchal institutions that enjoy great social control; second, that the feminist case for pornography as group libel comes out of a history of sexism; and third, that sexism is not a superficial product of public life, but rather an internalized element of identity as something partly constituted by gender.

I have now provided a definition of group libel and made explicit its forms. I have established my task as proving that pornography undermines the dignity of women, and I have made my assumptions known. I will now proceed with my argument.

Pornography undermines the dignity of women by disseminating misogynist ideals. The ideals are neither

inconsequential nor arbitrary. First I must clarify what I mean by “misogynist ideals”. I call “misogynist ideals” those ideas about social worth that are generated and maintained by patriarchal control. On this view, prevailing ideas or themes in pornography are misogynist ideals insofar as they are created by and for men.

One central misogynist ideal in pornography is the representation of women as essentially sexual. Catherine MacKinnon argues in the journal *Signs* that the female stereotype is fundamentally sexual. It is constructed so that “femaleness means femininity, which means attractiveness to men, which means sexual attractiveness, which means sexual availability on male terms”.<sup>30</sup> The patriarchal construction of female stereotypes is possible because gender is a learned quality. It is “an acquired characteristic, an assigned status, with qualities that vary independent of biology and an ideology that attributes them to nature”.<sup>31</sup> Women who fail to emulate the idea of the woman as “docile, soft, passive, nurturing, vulnerable, weak, narcissistic, childlike, incompetent, masochistic, and domestic” are thought of as “less female, lesser women”.<sup>32</sup>

The sexualized female stereotype is an ideal that lies at the heart of pornographic material. Pornography depends on the stereotype of women as fundamentally sexual by definition. As an industry “that mass produces sexual intrusion on, access to, possession and

use of women by and for men for profit”, pornography functions as pornography only when it portrays women in this light.<sup>33</sup> This argument does not depend on blurring the differences in details across a variety of pornographic films. All pornography that features women, features women as sexual commodities for men regardless of its particular content. “The most pernicious message relayed by pornography,” argues Adrienne Rich, “is that women are natural sexual prey to men and love it; that sexuality and violence are congruent; and that for women sex is essentially masochistic, humiliation pleasurable, physical abuse erotic”.<sup>34</sup> Even lesbian and “soft-core” pornography reinforce these ideals. Rich points out that “sensuality between women ... is ... pornographic in itself” because the absence of a male sexual partner is queer, different, and erotic; and soft-core pornography depicts women “as objects of sexual appetite devoid of emotional context, without individual meaning or personality”.<sup>35</sup>

Especially troubling is the content of “hard-core” or violent pornography. This type of pornography is uniquely troubling for two reasons. It is troubling, first, because what men masturbate to in these films is “women being exposed, humiliated, degraded, mutilated, dismembered, bound, gagged, tortured, and killed”.<sup>36</sup> It is troubling, second, because the production and consumption of such pornographic material desensitizes audiences to sexual violence as violence. As “masturbation material”, women’s

suffering becomes eroticized.<sup>37</sup> As MacKinnon puts it, “your violation [becomes] his arousal, your torture his pleasure”.<sup>38</sup> An excerpt from the 1983 Minneapolis Ordinance reinforces the point. The ordinance, drafted by Catherine MacKinnon and Andrea Dworkin, claims that “pornography is the sexually explicit subordination of women, graphically depicted, whether in pictures or in words”. It notes that in pornography “women are presented as dehumanized sexual objects, things or commodities; ... as sexual objects who enjoy pain or humiliation; ... as sexual objects who experience sexual pleasure in being raped”.<sup>39</sup> One section asserts that “women are presented as whores by nature”.<sup>40</sup> The ordinance continues on in this vein.

A second misogynist ideal in pornography is women’s effortless ability (or inability) to separate sexual desires from emotional feelings. In her essay *Exploring the Managed Heart*, Arlie R. Hochschild evaluates the role emotional services play in the work of customer service employees like flight attendants. She finds that “the emotional style of offering the service is part of the [flight attendant’s] service itself, in a way that loving or hating wallpaper is not a part of producing wallpaper”.<sup>41</sup> In addition to physical and mental labour, says Hochschild, the flight attendant is expected to perform emotional labour. Emotional labour “requires one to induce or suppress feeling in order to sustain the outward countenance that produces the proper state of mind in others—in this case, the

sense of being cared for in a convivial and safe place”.<sup>42</sup> In other words, the flight attendant is expected, first, to physically carry out her job; second, to engage mentally with the work she is doing; and third, to suppress her emotions so that they do not interfere with the desired emotions of the paying customer. Hochschild observes that half of the female workforce consists of jobs that call for emotional labour. More women than men have emotional responsibilities at work, she notes, and “they know more about its personal costs”<sup>43</sup>. The phenomenon is the commercialization of emotions as a necessary precondition of economic survival as a woman.

I think nowhere is the commercialization of emotion more prominent than in the pornography industry. Hochschild points out that airlines sexualize their flight attendants with slogans like “We really move our tails for you to make your very wish come true” and “Fly me, you’ll like it”.<sup>44</sup> This sexualisation, she says, fulfils some male fantasy of the flirty and obedient lady. Pornography obviously fulfils the male fantasy, too. But while flight attendants are expected to feign an emotional connection with their customers, porn actresses are expected to be able to feign an emotional connection in certain instances, and sexual attraction devoid of emotional content in others.

Pornography thus depicts the misogynist ideal that women do not have lasting or legitimate feelings unless it is sexually satisfying for men. Women are not only sex

objects, but they are customizable sex objects: they can be emotionally turned on or off depending on the sexual need of the paying customer. The danger here is that, as MacKinnon argues, “as society becomes saturated with pornography, what makes for sexual arousal, and the nature of sex itself ... change”.<sup>45</sup> Pornography is not mere “representation”, where “representation” means non-reality. What is experienced in pornography is real; the emotional emptiness expected of actresses in pornography translates to how we understand sex in the real world.

Flight attendants, too, are surely expected to exercise flexibility in their emotions. The difference is that the rewards of their emotional labour—say, from the successful execution of a great in-flight service—are somewhat tangible. Flight attendants interact with their customers; polite customers thank them for their service, and the quality of their service affects their career mobility. Porn actresses are simply consumed: their immortalization in film as sexual objects brings no such rewards. Every pornographic film asserts something about her sexuality, about the “nature” of her gender as sexual, as emotionally void, as object.

Pornography thus understands woman as such: she is essentially a sexual being available for sexual access; she is naturally sexually subordinate to men and enjoys her subordination; she finds sexual pleasure in violence committed against her (and welcomes it); and she is a sexual object, devoid of human emotion

and personality. This is how all women, regardless of their particularities, are depicted in pornography. MacKinnon puts it like this: “pornography says that women are a lower form of human life defined by their availability for sexual use”.<sup>46</sup> I turn now to showing the ways in which pornography qualifies as group libel.

Pornography is first an instance of libel by factual claim. I argued the claims that women enjoy sexual violence and sexual subordination are expressed in all types of pornography involving women. The claims are “factual” insofar as they are expressed universally in pornographic materials. Distributing pornography therefore always amounts to distributing false claims about women as a group. The effect is that any individual woman—in virtue of her group membership as a woman—is denied the dignity and respect society owes her. Her dignity is undermined because she is represented as a sexual object in pornographic media, and remains a sexual object when she is cat-called on the street or sexually assaulted.

She may not feel her status as a citizen in society has been diminished; she may even genuinely enjoy the attention. As I pointed out, however, group libel is not concerned with subjective feelings; group libel is concerned with how things are. A woman need not feel like a sexual object to be regarded as one by men; she need not feel subordinated to have her identity as a woman developed and forced upon her as something fundamentally sexual. Rae Langton quotes

Easterbrook: “Depictions of subordination ... tend to perpetuate subordination. The subordinate status of women in turn leads to affront and lower pay at work, insult and injury at home, battery and rape on the streets”.<sup>47</sup>

An objection is to say that some women do indeed enjoy sexual violence and that what is depicted in pornography does not harm them because they agree, wholeheartedly, with the ideals of the pornographer. I believe objections of this sort falsely equivocate the realities of sexual violence with the potential pleasures of “rough sex”. A woman cannot enjoy sexual violence because sexual violence is suffering, suffering is harm, and harm is unwanted. It is an oxymoron to suggest that someone wills their own harm, or that they enjoy it. The very act of willing one’s own harm contradicts the definition of harm, as does the act of enjoying it. The point I am making is not that masochism is impossible, but rather that “sexual violence” as I have used it here refers exclusively to unwanted and painfully endured sexual behaviours. This objection dissolves when this distinction is made clear.

Pornography also qualifies as libel by denigrating claim. Not only does pornography assert false factual claims about women (for instance, that women enjoy sexual violence), but it also asserts derogatory opinions that undermine women’s social and cultural reputation. Keegstra sought to convince his students that Jewish people were malicious and calculated liars.

The pornographer seeks to convince his audience that women are passive and unfeeling sexual objects, readily available for sexual consumption at the whim of men. He does this by paying her to enjoy sexual violence, degradation, and the consumption of her body as a mere sexual object. He also does this by making erotic traditional women’s roles: he films scenes where secretaries give in to their bosses’ sexual advances to revive the historical stereotype that “secretaries are for sex”. The pornographer thus shapes the social image and reputation of women; his declaration that “secretaries are slutty” manipulates workplace environments and workplace relationships. He models all women after misogynist ideals and sells their denigrated image for profit. The effect is that oppressive gender roles become deeply entrenched in our social fabric.

Depicting women as sexual objects is also libel by dehumanizing claim. The vicious characterization of the female body as a mere sexual object dehumanizes the entire group at the expense of the individual. It dehumanizes women by equating the sum of the sexual appeal of their body parts with their total worth as women. If to be a woman is to be a sexual object, excellence in womanhood is synonymous with sexual appeal for men. What is dehumanizing about this process is different from what made calling African Britons “apes” dehumanizing. In racist discourse, libel by dehumanizing claim tends to equate some group with animals or insects. In sexist discourse, this

sort of libel is ineffective for the misogynistic task at hand. Unless comparisons with insects or animals reassert women's subordinate position as a sexy thing, dehumanizing a woman consists in stripping the woman of her subjectivity and leaving behind only her body. The belief that "all humans, whatever their color or appearance, are equally persons with the rights and dignity of humanity" comes to include only men.

Pornography is libel by discriminatory claim as well. As I mentioned, any sign or emblem like "Muslims Out!" that intends to signal the degradation of a group is an instance of libel by discriminatory claim. Pornography sends no such explicit messages: "Women are less than men!" is hardly a selling point for the pornographer. Rather, the libellous emblems of pornography are those that describe women using the misogynist language of sexual objectification. They are the signs that hang outside of theatres and strip clubs that advertise "hot women" and "live sex shows". Signs of this kind demand that women be objectified, thought of as lesser citizens of society, and made other. The sign that advertises big breasted women at a strip club advertises them as sexual objects. The sign also demands, in virtue of having done it itself, that its viewers objectify all women on the basis of their physical attributes. The social environment is poisoned insofar as women's participation in it becomes possible only on male terms, terms that subordinate women as a group. Waldron adds, "not only does pornography

present itself as undermining society's assurance to women of equal respect and equal citizenship, but it does so effectively by intimating that this is how men are taught, around here, on the streets and on the screen, if not in school, about how women are to be treated".<sup>48</sup>

In sum, I have presented Jeremy Waldron's framework of hate speech and argued for pornography as an instance of group libel. My motivation was to attract legal attention to this industry that profits by subordinating and encouraging the subordination of women. Pornography is an interesting case study in hate speech. It requires that we first address endemic sexism as the possibility for its being hate speech, and second that we identify its deeply embedded sexist ideas. In a society that continuously rehearses sexist ideals in its everyday, the task is not easy. I have attempted to show the ways in which women's dignity and assurance are undermined by pornography; its objective to objectify women, its eroticization of violence, and its invalidation of women as independent, real, feeling persons were among them.

In conclusion, I assert the urgency of the matter at hand. If my arguments hold and pornography is group libel, then it is the libel of an entire gender. Pornography must be understood by the law as a mechanism of sexist control. Only then will its harmful ideas and misogynist ideals be taken seriously as the real oppression of real women. The matter is urgent because pornographic

ideals shape our understanding of gender and self. Women will never be equal if society is taught that they are objects; they will never share in the dignity afforded to men if they are objectified by everyone—even their fellow women—as sexual commodities. Liberating women from sexist oppression depends on reshaping our understanding of “women” as a group. The regulation of pornography—understood not as “art”, or “freedom of expression”, or “just images”, but as group libel—is a great step towards this goal.

### **Works Cited:**

- 1 Waldron, Jeremy. “Approaching Hate Speech,” *Harm in Hate Speech*, 2012. p. 4.
- 2 Ibid. p. 4.
- 3 Ibid. p. 4.
- 4 Ibid. p. 11.
- 5 Ibid. p. 45.
- 6 Ibid. p. 66.
- 7 Ibid. p. 3.
- 8 Ibid. pp. 39-40.
- 9 Ibid. p. 47.
- 10 Ibid. p. 56.
- 11 Ibid. pp. 56-57.
- 12 Ibid. p. 57.
- 13 Ibid. p. 57.
- 14 Ibid. p. 48.
- 15 Ibid. p. 58.
- 16 Ibid. p. 58.
- 17 Ibid. p. 58.
- 18 Ibid. p. 59.
- 19 Ibid. p. 59.
- 20 Ibid. p. 59.
- 21 Ibid. p. 60.
- 22 Ibid. p. 106.
- 23 Ibid. pp. 82-83.
- 24 Ibid. p. 87.
- 25 Rich, Adrienne. “Compulsory Heterosexuality and Lesbian Existence,” *Women: Sex and Sexuality*, 1980, 62-91. p. 633.

- 26 Scott, Joan. "Experience," *Feminists Theorize the Political*, 22-40. p. 30.
- 27 Bordo, Susan. "The Body and the Reproduction of Femininity," *Unbearable Weight: Feminism, Western Culture and the Body*, 1995, 165-184. p. 165.
- 28 Ibid. p. 166.
- 29 Ibid. p. 173.
- 30 MacKinnon, Catharine. "Feminism, Marxism, Method, and the State: An Agenda for Theory," *Signs*, 1982. p. 525.
- 31 Ibid. p. 529.
- 32 Ibid. p. 5.
- 33 MacKinnon, Catharine. "Pornography: On Morality and Politics," *Toward a Feminist Theory of the State*, 195-214. p. 195.
- 34 See Rich, p. 641.
- 35 Ibid. p. 641.
- 36 MacKinnon, Catharine. "Defamation and Discrimination," *Only Words*, 1993, 3-33. p. 17.
- 37 Ibid. p. 17.
- 38 Ibid. p. 4.
- 39 Duggan Lisa, Hunter Nan, Vance Carole. "False Promises: Feminist Anti-pornography Legislation," *Sex Wars: Sexual Dissent and Political Culture*, 2006, 43-63. p. 62.
- 40 Ibid. p. 62.
- 41 Hochschild, Arlie. *The Managed Heart: Commercialization of Human Feeling*. pp. 5-6.
- 42 Ibid. p. 7.
- 43 Ibid. p. 11.
- 44 Ibid. p. 93.

- 45 See MacKinnon, Catharine. "Defamation and Discrimination", p. 25.
- 46 See Waldron, p. 90.
- 47 Ibid. p. 94.
- 48 Ibid. p. 91.

## Intuitions in Philosophical Thought Experiments

Abel Girma

**Abstract:** Thought experiments are a fundamental tool in philosophy used to help tease out our intuitions on a host of issues. This paper questions the role of thought experiments in enabling us to make epistemically significant claims. While carefully constructed thought experiments appear to reveal our intuitions about knowledge, our intuitions do not always seem to respond to epistemically meaningful factors. The findings in this paper suggest that we ought to be more skeptical about the ways in which we use and rely on intuitions in thought experiments to develop theories of knowledge.

## I. Introduction

What does it mean for me to know that roses are red? A great number of philosophers have sought to answer this question and, more generally, to develop the best theory for understanding knowledge. Many have attempted to do so by constructing thought experiments and deriving conclusions from intuitive responses to them. Traditionally, knowledge has been defined as justified true belief (JTB). What it means for me to know that roses are red is for me to believe that roses are red, for me to have justification for that belief and for it to be true that roses are red. However, Edmund Gettier demonstrated that this theory is unsatisfactory—an agent can have a JTB but still fail to know.<sup>1</sup> “Gettier cases” are thought experiments in which elements such as luck or coincidence defeat an agent’s knowledge, despite their having a justified true belief.

Many philosophers have attempted to tweak JTB by, for instance, adding a fourth condition. For example, Lehrer and Paxson have—ultimately unsuccessfully—constructed variations of Gettier cases to argue that knowledge is an undefeated JTB.<sup>2</sup> Another such attempt has been to contend that knowledge is JTB without a false premise. Alvin Goldman showed through his Fake Barn Country thought experiment that no false premise as a condition to address the Gettier problem also fails.<sup>3</sup> No uncontested and thoroughly persuasive response to Gettier cases seems to exist. In this

ongoing debate, prominent theorists have argued that the debate surrounding the Gettier problem suggests that we ought to have a different approach in order to appropriately define knowledge.

While carefully constructed thought experiments appear to reveal our intuitions about knowledge, our intuitions do not always seem to respond to epistemically meaningful factors. Gendler and Hawthorne's discussion of the Fake Barn Country thought experiment<sup>4</sup> suggests that our intuitions in these cases are not reliable in the way we expect them to be. In particular, it seems that our intuitions can be influenced by epistemically insignificant information, making them an unreliable tool for developing an epistemically important conception of knowledge. If knowledge is to be defined as a concept that is epistemically important, then only epistemically significant factors should affect our judgment with respect to cases of knowledge. A consideration of intuitive responses to variations of the Fake Barn Country thought experiment shows that our intuitive responses might not be a reliable tool to help us develop an epistemically satisfactory theory of knowledge.

## II. First Case: Original Fake Barn Country

Henry, an ordinary person, is driving down a road in a rural town. Looking out the window, he begins to identify the various objects he sees. "That's a cow," "That's a tractor," "That's a Barn," says Henry. Henry

is especially confident that he sees a barn—he has an unobstructed view, great eyesight and enough time to observe the object and identify it as a barn. Moreover, what Henry sees and identifies is, in fact, a barn. Intuitively, it appears that Henry knows it is a barn.

Unbeknownst to Henry, however, the road he is driving on is lined with many fake barns—deceptive façades of barns. These fake barns are so deceptive that they are indistinguishable from real barns when seen from the road. In fact, the object that Henry identified as a barn happens to be the only real barn on the road. Does Henry know it is a barn? It seems unintuitive to claim that Henry knows that the object he is seeing is a barn. The fact that Henry looks at that particular barn is primarily due to luck.

Goldman provides a theoretical account for why Henry does not know. Despite Henry's having a true belief, Goldman claims, the presence of "relevant counterfactual situations," in which an equivalent perceptual experience would have led to a false belief defeats his knowledge.

## III. Second Case: One Fake Barn Country

Let us now imagine a slightly different story, in which Henry is driving down a different road. Again, Henry names the objects he sees: "That's a cow," "That's a tractor," "That's a Barn." In fact, Henry does correctly identify each object, including the barn, as in the first case. Intuitively, it seems that Henry knows it is a barn.

Unbeknownst to Henry, it turns out that there is a fake barn—a powerfully deceptive façade of a barn—further down the road. Unlike the original case in which there were “many” fake barns, however, there is only one fake barn on this road. If Henry drives for a bit longer, he would pass by the fake barn. Does Henry know that he is seeing a real barn? Intuitively, we do not seem to have a strong impulse to say Henry does not know (compared to our intuitive response in the original case). The mere presence of a single fake barn down the road does not seem like a strong enough reason for Henry’s knowledge to be defeated.

When the story told includes “many” fake barns, we readily determine that Henry does not know. When it is specified that there is only one fake barn down the road, the intuitive pressure to rid Henry of knowledge disappears. However, with one other fake barn on the road, Henry is still in significant danger of identifying the fake barn as a real one. The risk of falsity remains quite high. If the epistemically important factor that defeated Henry’s knowledge in the original case was the risk of falsity, the presence of one fake barn should have the same effect on our intuitions. However, our intuitions appear to be responding to an epistemically insignificant change.

One can argue that there is an epistemically significant difference between the presence of one as compared to “many” fake barns. With many fake barns, the risk of Henry’s being wrong is very high, which

intuitively warrants his knowledge to be defeated. The more fake barns there are, the higher the risk of falsity. Thus, it could be that the intuition that Henry knows in the one fake barn case might be reflecting the decreased risk of falsity.

Nevertheless, One Fake Barn Country puts pressure on how our intuitions respond to a tweak in a thought experiment, which does not seem epistemically very important. While “many” fake barns pose a higher risk of falsity than one fake barn, a fifty percent chance of being wrong is still very high. The extent to which our intuition about Henry’s knowledge changes between the two cases is not proportional to the extent to which epistemically important factors change between the two cases. In other words, while little seems to change with respect to epistemically important factors—namely, the risk of falsity—our intuition about Henry’s knowledge changes considerably.

#### **IV. Third Case: Fake barns in Barnywood**

Henry is driving to the airport to board a plane to a nearby town when he sees a barn by the road. As in the first case, Henry confidently says: “That’s a barn!” In addition, there are no other barns—real or fake—in the area. In the absence of deceptive barns nearby, it seems that Henry knows it is a real barn.

In half an hour, Henry lands in Barnywood, where the road leaving the airport is lined with fake barns. Henry rents a car and starts driving on that road.

Intuitively, it does not seem like the existence of fake barns in this nearby town affects Henry's knowledge that there is a barn in the first town. But what is the epistemically relevant difference between this thought experiment and the original Fake Barn Country case?

In the original Fake Barn Country case, our intuition that Henry does not know was due to the fact that Henry could have just as likely pointed to a fake barn and incorrectly identified it as a barn. In this thought experiment, Henry could have just as likely pointed to a fake barn in Barnywood and incorrectly identified it as a barn as well. In fact, we may add that in both the original Fake Barn Country case and the Barnywood case, Henry's entire trip takes 2 hours. In Fake Barn Country, we can imagine that Henry is driving for two hours on the road with fake barns. Similarly, we can that his drive to the airport, flight to Barnywood, and drive on the road in Barnywood take a combined 2 hours.

Thus, the risk of Henry incorrectly identifying a fake barn as a real barn does not seem to change in the two cases. It does not seem like anything relevant to determining whether or not Henry knows in the original case is affected in the second case. If our intuition in the original case is that Henry does not know it is a barn, it is worrisome that our intuition in this case is different, despite no apparent epistemically important change.

Again, one might object by arguing that our intuition is changing because the change in geographic location is epistemically relevant. While Henry's perceptual experiences are reliable on the first road, they are not reliable on the new road lined with fake barns. On the first road, there is no danger of falsity, while on the second road, Henry would be wrong to identify any of the façades as a barn. As such, it would seem reasonable for our intuition about Henry's knowledge to be different in the original Fake Barn Country case and the Barnywood case.

This response is insufficient. If we specify, as we have, that Henry's entire trip in the both cases takes two hours, then Henry is just as likely to have incorrectly identified a façade in both cases. If we establish the risk of falsity as the epistemically important factor that intuitively defeats Henry's knowledge in the original case, it should follow that the risk of falsity in the second case should affect our intuition in the same way. As Gendler and Hawthorne explain: "It seem[s] quite unclear how to calibrate live danger so that lines are drawn where intuition suggests they ought to fall."<sup>5</sup>

## **V. Fourth Case: Henry Stranded in Fake Barn Country<sup>6</sup>**

Let us now imagine that the barns in Fake Barn Country are distanced such that the risk of perceiving fake barns after seeing the singular real barn is very high if one is driving, but very low—essentially

nonexistent—if one is walking. That is, the fake barns are not within walking distance from where Henry is when he sees the real barn. Additionally, let us assume that the real barn is the first barn on the road from where Henry is driving.

Let us now suppose that Henry’s car breaks down immediately after he identifies the real barn. He is stranded on the road, with only the real barn in sight, and the nearest fake barn too far away to walk to. Intuitively, it seems like Henry knows he is seeing a real barn. In fact, now, there is no longer a “relevant counterfactual situation” in which Henry would have been false in his claim that what he sees is a barn.

To see how this is especially troubling, let us imagine that Henry is actually a car mechanic and can quickly fix his car. Once he fixes his car, Henry can continue driving on the road, where he will soon pass by numerous fake barns. All of a sudden, the risk of Henry perceiving and identifying fake barns seems to change our intuition that Henry knows he is seeing a real barn. Surely Henry’s auto-mechanical expertise should have no epistemic weight in determining whether or not he knows that what he is seeing is a real barn. Still, it seems that our intuition about whether or not Henry knows can depend exclusively on whether or not he is an auto-mechanic when his car breaks down.

Proponents of Goldman’s analysis of the Fake Barn Country thought experiment might respond by arguing that something epistemically important is changing

in this story. Henry’s profession is not the significant factor in determining whether or not he knows. What is changing our intuition about Henry’s knowledge claim is not his job title but rather the presence or absence of a relevant counterfactual situation in which he would be wrong. If Henry is a car mechanic, then nothing important about the original Fake Barn Country case changes when his car breaks down. He can easily repair it and face the same issue discussed in the original case again. If Henry does not have the ability to fix his car, on the other hand, the risk of falsity and the relevant counterfactual situation disappear.

Nevertheless, this explanation seems too ad hoc. To appeal to the change in relevant counterfactual situations as the important factor changing our intuitions presupposes Goldman’s theory as the appropriate conception of knowledge. The explanation does not address the main issue we are attempting to highlight: that our intuitions are easily swayed by details that should have no direct import to the epistemic question at hand.

## **VI. Fifth Case: Henry with nearsighted Maya<sup>7</sup>**

In this new thought experiment, let us introduce Maya, Henry’s friend, who is in the car with Henry. They are driving down a road with one real barn and no deceptive fake barns. Henry and Maya both point to various objects, as in the original Fake Barn Country case, and both identify the singular real barn. Nothing

seems to interfere with their knowledge so far. In addition, let us suppose that Maya is nearsighted; she can only see a few dozen yards beyond the road on either side. So, while she can see the barn very clearly, she cannot see much further past the barn.

On the other side of the road a river flows, bordered by another road—Riverway Drive. Unlike the road Henry and Maya are driving on, Riverway Drive is lined with many fake barns. However, while these fake barns are too far for Maya to see, they are as perceptually accessible to Henry as the real barn is. Following our intuition from the original case, Henry does not seem to know that he is seeing a real barn when he identifies the object as a real barn. This is because the fake barns on the other side of the river make it so that Henry could have just as likely pointed to a fake barn and incorrectly identified it as a real barn. Their presence creates relevant counterfactuals that block Henry's knowledge. Of course, Maya, on the other hand, could not have made a similar mistake. For Maya, no relevant counterfactuals exist because she could not have seen the fake barns given her nearsightedness.

The problem that arises should be clear: how much better Henry's vision is should not matter in determining who knows and who does not know as long as they are both able to clearly see the real barn. Once again, it seems like our intuition with respect to knowledge attribution is being determined by epistemically insignificant factors.

## VII. Conclusion

Gendler and Hawthorne offer a humorous synthesis of the flaw in our intuitions with respect to the thought experiments discussed:

The walker knows; the driver doesn't. The short-sighted observer knows; the observer with 20-20 vision doesn't ... It is intolerable, they argued, to allow that pedestrianism and shortsightedness could yield epistemic dividends in this way. Perceptual risk, they maintained, is highly observer-sensitive - in ways that knowledge is not - so the two cannot go hand-in-hand.<sup>8</sup>

More importantly, in all the cases discussed above, it seems like our intuitions are highly influenced by epistemically insignificant—or at least not considerably significant—information in a way that knowledge should not be. Why then should we rely on our intuitions to give us meaningful guidance in defining knowledge?

Laurence Bonjour argues that any correct conception of knowledge should reflect the supreme value we attribute to knowledge.<sup>9</sup> He asserts that only absolute certainty can count as knowledge. In other words, the level of epistemic justification needed to achieve knowledge must be such that it guarantees knowledge. On his account, Henry would not know it is a barn in any of the cases because he would not be certain it is a barn. Bonjour's explanation of knowledge has been largely received as unintuitive, while less demanding

conceptions of knowledge have been given much more attention, in large part due to their intuitive appeal.

The findings in this paper suggest that we ought to be more skeptical about the ways in which we use and rely on intuitions in thought experiments to develop theories of knowledge. As such, it reveals the importance of considering non-traditional theories of knowledge when formulating the epistemology of the future.

### **Works Cited:**

- 1 Edmund Gettier, “Is Justified True Belief Knowledge?” *Analysis*, vol., 23 (1963), 21–23.
- 2 Lehrer, Keith, and Thomas Paxson. “Knowledge: Undefeated Justified True Belief”. *The Journal of Philosophy* 66.8 (1969): 225–237.
- 3 Goldman, Alvin I. “Discrimination and Perceptual Knowledge”. *The Journal of Philosophy* 73.20 (1976): 771–791.
- 4 Gendler, Tamar Szabó, and John Hawthorne. “The Real Guide to Fake Barns: A Catalogue of Gifts for Your Epistemic Enemies”. *Philosophical Studies: An International Journal for Philosophy in the Analytic Tradition* 124.3 (2005): 331–352.
- 5 See Gendler & Hawthorne, p. 338.
- 6 Adapted from brief discussion of Ed and Fred in Gendler & Hawthorne pp. 338-339.
- 7 Adapted from “DAYTIME VOYAGE” case in Gendler & Hawthorne, p. 339.
- 8 See Gendler & Hawthorne, p. 339.
- 9 BonJour, Laurence. «The Myth of Knowledge.» *Philosophical Perspectives* 24.1 (2010): 57-83.

## Pleasure and the Nature of Complete Activity

Cait Duggan

**Abstract:** In this paper I will investigate two points: first, how activities that are processes (e.g. building) can be understood as pleasurable, and second, how exactly pleasure completes activity, which will sharpen our understanding of the relationship between processes and pleasures. Aristotle states that pleasures arise from activities that exhibit their form in a moment and are not evolving towards it. I will argue that processes, which Aristotle says cannot be sources of pleasure, can be viewed as pleasurable on account of the fact that their forms have some connection to the end states they strive towards. I will also argue that we can tie the concepts of pleasure and processes closer together by noting how both are like the activity of contemplation.

In X.4 of the *Nicomachean Ethics*, Aristotle uses the intriguing phrase “the nature of complete activity”<sup>1</sup> to describe the kind of activity during which people experience the most pleasure. Pleasure can arise from this kind of activity but is not itself a process, the form of which requires time to manifest itself in full.<sup>2</sup> When we experience pleasure, the form of pleasure is present in an instant; there is no difference between the beginning and the end of our experience of pleasure, as there is between the beginning and the end of a process. This account of complete activity will bring us a great deal of confusion when we recognize that it renders a process as somehow less of an activity than “complete activities,” and thus not as pleasurable as they. In this paper I will investigate two points: first, how activities such as building can be understood to be pleasurable once we refer to Aristotle’s *Metaphysics* Theta 6. This should assuage some worries we might have about the plausibility of Aristotle’s account, since saying that pleasures are not processes runs contrary to our experience of various processes as pleasurable, such as the processes of building, painting a picture, growing plants, or cooking. This first question leads to a second inquiry, namely, how exactly pleasure completes activity, and how this adds to our understanding of the relationship between processes to pleasures. This will require a discussion of the different ways in which Aristotle uses the term “completion.”

Aristotle offers us a negative and a positive characterization of pleasure. Providing us the negative one first, he says, “... [pleasure] is not a process either, since every process—building, for example—takes time and has an end, and is complete when it has produced what it aims at, it is complete, therefore, in the whole time it takes, or at the final moment”.<sup>3</sup> Processes necessarily evolve towards something. The stages in a process differ in form, and their form is only manifested in their final moment when they have reached their end, whereas the entire form of pleasure exists in a moment. When positively characterizing pleasure, Aristotle writes, “but the form of pleasure is complete at any given moment, so it is clear that it is different from a process, and that pleasure is something whole and complete”.<sup>4</sup> Thus, Aristotle makes a distinction between the form of pleasure and the form of a process, which casts doubt on the possibility of pleasure being a process. Furthermore, at the conclusion of X.4 he notes that pleasures can only arise from activity and are specific to the activities from which they come.<sup>5</sup> What is crucial about pleasures is that they arise from activities that exhibit their form in a moment, and are not evolving towards it.

Aristotle also provides us with a few words on the nature of complete activity. He offers a contextual definition of complete activities as those in which, “every sense engages in activity in relation to its object, and its activity is complete when it is in good condition

in relation to the noblest of its objects”.<sup>6</sup> Thus, we see that senses have objects that are suited to them specifically. Seeing is Aristotle’s paradigmatic example of a complete activity, because the form of the activity of seeing is present in its fullness in the moment in which we are seeing something. In other words, unlike building, there is no further end that the activity of seeing is working towards in the moment when I am seeing something, other than my seeing the object I am looking at. We could certainly be looking at something, such as a page, for the purpose of reading the words on it and acquiring some knowledge, but this acquiring of knowledge is distinct from seeing. After describing complete activity, Aristotle articulates the relationship between complete activity and pleasure when he writes, “When both object and subject of perception are at their best, there will always be pleasure, since what will produce it and what will experience it are both present”.<sup>7</sup> Aristotle does not say how difficult it is for us to get our senses to operate at their best, or how difficult it is to fix our attention upon the noblest of objects, but he later says that it is impossible for us to be engaged in complete activity constantly, indicating that we find ourselves busy with this special kind of activity in an inconsistent way.<sup>8</sup>

The initial question as to how we can make sense of the pleasure we derive from processes is a question about whether we can derive pleasure from activities that are not complete, even if these are not the most

pleasant undertakings. If we look to Aristotle's *Metaphysics* Theta 6, there is an explanation of processes in which a process's end is present in it in a potential way. Aristotle writes:

Actuality means the existence of the thing, not in the way which we express by 'potentially'; we say that potentially, for instance, a statue of Hermes is in the block of wood and the half-line is in the whole, because it might be separated out, and even the man who is not studying we call a man of science, if he is capable of studying.<sup>9</sup>

Here we see that the end towards which something is moving exists in that object as it moves towards that final point. This is an end that exists in an object before the process that is working on this object is brought to completion in its final moment, presumably by the agent carrying out the process. Aristotle says that processes are made of parts that differ in form,<sup>10</sup> but while there is certainly a difference between "placing stones together and fluting a column, and these are both different from the production of the temple",<sup>11</sup> these can both be viewed as steps in the more general process of building a temple. The successful production of the temple Aristotle says is complete, but the process of building the temple is not. Thus, in that sense, one could say that these activities share the form of the process of building a temple, in addition to their own respective processes of collecting stones in a single

place and cutting stone. The early stages of a process can be viewed as pale shades of the form of the final product, which would be the production of the temple.

This passage from the *Metaphysics* acknowledges that objects that are being worked on in processes are in states of potentiality with respect to their final state of completion, even though it does not give us the grounds to say that pleasure is a process. It does, however, provide a more optimistic account of how objects that are in states of potentiality can be an image of their completed end. This view from the *Metaphysics* approaches processes from a different angle than the angle from which they are discussed in the *Nicomachean Ethics* by emphasizing not their incompleteness, but rather their proximity to their end. The discussion of potentiality here is not meant to conclude that pleasure is now a process; rather it is only meant to suggest a positive account of how, despite the fact that processes are evolving towards their forms, they could be pleasurable because their forms have some connection to the end states that they strive towards.

One commentator, David Bostock, provides an explanation of how processes can be pleasurable in his paper, "Pleasure and Activity in Aristotle's *Ethics*," that supports the view I have proposed here. In his argument he discusses whether restorative pleasures originate in the body or in the mind, but I will focus on his discussions of the idea of restoration. He uses

Aristotle's discussion of restorative pleasures in Book VII to argue that the pleasure from processes is like the pleasure of replenishment.<sup>12</sup> He notes that Aristotle believes that the pleasures that we derive from restorations are pleasures that arise from the part of us that is not depleted, the part that is not in need of restoration.<sup>13</sup> Rather, the pleasures arise from the perception of this restored and healthy part of us.<sup>14</sup> If we are hungry, this could simply be the perception of being full, of being hungry no longer. The suggestion I have made is that we can experience pleasure when we are engaged in a process because the process still has a connection to its end, which can be viewed on the restorative pleasure model as the part of us that is not lacking, the part that is completed. There is no other end in mind that our process could have other than one towards which it is working. Stages of potentiality towards an end can be viewed as lacks, such as hunger or thirst. Just as Aristotle thinks that we find incidental pleasure in restorative pleasures from the part of us that is not lacking, so we find pleasure in processes coming from the image of the end that is inherent in the process and that which gives the process its form. A process's motion towards the completion of its form should be viewed on the model of restoration.

Thus, if activities that are processes do contain their end, in a mitigated way, then perhaps we can then say that these activities can actually be pleasurable. Still, one could reasonably object that the half completed

building is very different from the finished building insofar as the latter has attained the end of building, while the other one has not. One is still complete and the other is not. At this point, it is important to note a particular feature of Aristotle's teleology, namely, its objectivity. Aristotle's discussion of pleasure reveals that there are at least three senses of completion with respect to activities. One sense can be easily understood because it is the goal that we consciously aim towards: the end of the activity of building is to have a final build house. The other understanding of completion that Aristotle has provided is that in which the senses are operating well and have taken on the best objects for themselves. However, the next sense of completion is at a higher level, which is more in terms of the activities that Aristotle thinks we are meant to undertake in our lives, whether or not we want to perform them. Aristotle asserts that there is a chief good, and that this will be the most complete, that is "complete without qualification" and is something we pursue for its own sake.<sup>15</sup> Since we aim at ends when we perform actions, we see a hierarchy of ends, and therefore a hierarchy of activities, taking shape based upon those that are chosen for the sake of other things, which are at the lower end, and that which is chosen for the sake of nothing else, at the top. Ends can thus be compared with one another and we can judge the completeness of an end based upon whether or not we choose it for its own sake. Thus, the activities that we undertake, be it

building or seeing—those that are processes and those that are not—fit into a hierarchy of activities and ends. Aristotle says that there is only one activity that we undertake for its own sake, and that is contemplation, “since nothing results from it apart from the fact that one has contemplated.”<sup>16</sup> While he does not provide a simple definition of contemplation, he does say that it is “the life in accordance with intellect.”<sup>17</sup> It is the activity of the exercise of knowledge, rather than the pursuit of it. Thus, if we are going to give a thorough assessment of processes and activities in terms of completeness, we need to broaden our horizons and not only compare them to one another, but also judge them in light of the vast array of activities we can undertake.

Thus far I have suggested that processes can be viewed as aiming at their end, which allows us to view them as a version of complete activity, or at least not completely unrelated to it. This understanding of a process brings the notion of process and complete activity closer together, which in turn can bring the notion of process and pleasure closer together. This can make it possible to see how on Aristotle’s view, processes can give rise to pleasure because they are shades or versions of the end that they are approximating. I want to draw another connection between processes and complete activity in this second part of the paper so as to argue at greater length that the concepts of process and pleasure are not incompatible. I hold that there is a certain closeness between processes and pleasure, and thus processes

and completion, which can be illustrated through their similarities to the activity of contemplation. Processes require some activity of the intellect that is like the activity of contemplation. Pleasure completes activities in a way that makes them like contemplation, namely, by making the end of an activity less of our focus.

For Aristotle, the activity of the gods was contemplation, which he considered to be the greatest activity that we could undertake. Contemplation is the complete activity that combines both the finest operation of the highest element in us, our intellect, with the objects that are “the highest objects of knowledge.”<sup>18</sup> It should be noted that this definition of contemplation lines up quite closely with the definition of complete activity, and thus the most pleasurable activity, provided earlier in X.4. Indeed, Aristotle writes, “... for a human being, therefore, the life in accordance with intellect is best and pleasantest, since this, more than anything else, constitutes humanity.”<sup>19</sup> Verity Harte supports this assessment of contemplation and pleasure in her article, “The Nicomachean Ethics on Pleasure,” where she says that, “For Aristotle ... human experience of the most valuable pleasures – like eudaimonia in general, at least according to NE x 7 – is a respect in which humans exceed their animal nature and share in the life of gods.”<sup>20</sup> Since contemplation is the activity that “constitutes” our humanity, it seems right to say that we should be trying to practice as much contemplation as we can. This leads us to regard all of

our other activities in light of this greatest instance of activity.

If we look at the process of building closely, we can notice that it is composed of parts, and that those parts require the operation of the intellect. Building requires that we make designs and blueprints of the thing we want to construct, that we determine and obtain the supplies that we want, and that we choose the best methods of construction. I am not claiming that within each process there is an experience of contemplation, that is, an experience of full intellectual activity. This would be a mistake, since many processes need external goods, while contemplation is self-sufficient. Furthermore, our intellect is likely operating at a higher level with better objects of thought when we are contemplating than when we are building. However, the activity of the intellect is something that is crucial to contemplation and to processes that require activity of the intellect, such as building. The more we uncover Aristotle's picture of contemplation, the more we see that this is the activity of thinking about thinking itself. Building is certainly not a process of thinking about thinking, but there is a similarity between what we do when we undertake processes, and what we do when we contemplate.

Looking now towards pleasure and complete activity, Aristotle does provide us with a description of what it is that pleasures do that makes activities less like processes. He writes, "And since activities are made

more rigorous, longer lasting and better by their proper pleasure, and impaired by foreign pleasures, it is clear that the two kinds of pleasure are very different."<sup>21</sup> The comments about foreign pleasures aside, pleasure has the ability to put the end that we are working towards out of view, the end that, once reached, will bring the activity to an end. Pleasure sustains an activity and makes us want to continue doing it. It makes us want to rest in the activity rather than hurry to the end, and in that way it hides or submerges the end from view. Similarly, Aristotle writes that contemplation is, "also the most continuous, since we can contemplate more continuously than we can do anything."<sup>22</sup> It, like pleasure, has a quality of timelessness about it. Like pleasure, it also puts the end out of view insofar as there is no end that is being worked towards. Aristotle notes that contemplation has to be liked for its own sake, since "nothing results from it apart from the fact that one has contemplated."<sup>23</sup> Thus in a certain way pleasure makes our activity more like the activity of contemplation, because pleasure submerges the end we are pursuing from our view and in doing so, makes it seem as though there is nothing beyond or ulterior to our present activity, which is precisely our experience when engaging in contemplation. In addition, it is a feature of pleasure that is over and above activity insofar as it completes activities that are already technically complete, in the first sense of "complete" in that the senses are functioning well and have the

noblest objects as their own. Aristotle remarks that, “Pleasure completes the activity not as the inherent state does, but as a sort of supervenient end, like the bloom on the faces of young men.”<sup>24</sup> Contemplation seems to be the same way with respect to our human activities insofar as it is this activity that is both proper to us as human beings and something that is in a way superior to human life.

An objector might counter that my attempt to derive an explanation of how processes can be pleasurable by resorting to a more abstract hierarchy of activities and our capacity for intellectual activity, would compel me to say that Aristotle would claim that we are always engaged in activity, since we are frequently exercising our intellects in some way. This would be a problem since Aristotle explicitly denies that we are capable of continuous activity.<sup>25</sup> However, this objection can be countered when we recognize that another feature of the perfectly complete activity that is contemplation is that it is self-sufficient. Building, although it might be an exercise of our rational capacities, is certainly not a self-sufficient or self-contained activity, since it requires building materials, tools, and space to build, to name a few requirements. Thus, we can be engaged in processes where we exercise our rational capacities, without having then to claim that we are constantly in a state of complete activity.

In X.7 of the *Nicomachean Ethics* Aristotle offers some peculiar remarks about the activity and the end

that is most proper to us. He says that the greatest life for us is both the exercise of our intellectual capacity, and therefore something that seems very proper to us as human beings, but Aristotle also writes of the life of contemplation that, “such a life is superior to one that is simply human, because someone lives thus, not in so far as he is a human being, but in so far as there is some divine element within him.”<sup>26</sup> One might be given to think that the human end would be something particular to human beings, but Aristotle claims otherwise. Our lives have a goal that is beyond the ends that we choose to aim at; even complete activities, such as seeing, are only relative ends in a hierarchy of ends and activities that, as human beings, we are meant to recognize and respect. This claim of Aristotle’s furthers the idea that his teleology is not entirely relativized to human beings’ immediate interests. It includes the idea that we perform activities and aim at ends that are distant from even our own human natures. Although we can say that some ends of our activities pertain more to our immediate aims, and that we often take the ends of those activities to be the satisfaction of those goals we have, such as completing a building or a painting, we can also see how this hierarchy that our activities are a part of gives them a certain value based on what is good for us as human beings, not just on what we want. With this hierarchy, Aristotle can be defended from the charge that all processes are incomplete activities and lack pleasure, since they

require activities that participate in the most complete and pleasurable activity we as human beings could perform, namely, the exercise of our intellects. Their completion and pleasure are not derived from the ends that we take them to have, but from ends that exist at a more abstract level.

Much has been left unsaid in this account of Aristotle's view of pleasure, processes, and complete activity, perhaps most importantly an account of the objectivity of Aristotle's teleology and what we desire as the ends of our actions versus what is good for us as human beings, regardless of our desires. If, through a study of Aristotle's teleology, we can demonstrate that Aristotle thinks that processes have more ends than solely the ends we desire when we engage in those processes, then perhaps we will find a way of viewing processes as more similar to complete activities than we initially thought. Still, I hope to have illustrated how processes and pleasure are not so at odds with one another as it may first seem in the *Nicomachean Ethics*. Processes bear the image of their ends in them, putting them in contact with their end and with the completeness necessary for pleasure to arise. Similarly, processes require the operation of the intellect, which is precisely the operation of the most complete and most pleasurable activity we could undertake. Pleasure does not complete activity in one sense in that it does not sharpen the senses to perform to the best of their abilities, or put them in touch with the noblest objects;

yet, it does make activities more like the greater end of our lives, contemplation. Pleasure will follow on the completion in the first sense if the senses of the agent are at their best and are engaged with the best objects for them,<sup>27</sup> but given that processes employ our rational faculties, it seems possible that we can also derive some pleasure from processes because in undertaking them we exercise the faculty that is the most suitable for us to exercise.

**Works Cited:**

- 1 Aristotle, *Nicomachean Ethics* (Cambridge: Cambridge University Press, 2000), X.4 1174b20.
- 2 EN X.4 1174a22
- 3 EN X.4 1174a22-25
- 4 EN X.4 1174b6-9
- 5 EN X.4 1175a20-1
- 6 EN X.4 1174b18-20
- 7 EN X.4 1175a36-37
- 8 EN X.4 1175a46
- 9 Aristotle, *Metaphysics Theta 6* 1048a30-35
- 10 EN X.4 1174a26-27
- 11 EN X.4 1174a28-30
- 12 David Bostock, "Pleasure and Activity in Aristotle's *Ethics*" *Phronesis* Vol. 33, No. 3 (1988), 267-9.
- 13 Bostock, "Pleasure and Activity in Aristotle's *Ethics*," 267
- 14 Bostock, "Pleasure and Activity in Aristotle's *Ethics*," 269
- 15 EN I.7 1097a39-41
- 16 EN X.7 1177b1-2
- 17 EN X.7 1178b6
- 18 EN X.7 1177a22-23
- 19 EN X.7 1178a5-7
- 20 Verity Harte, "The *Nicomachean Ethics* on Pleasure" in R. Polansky, ed., *The Cambridge Companion to Aristotle's Nicomachean Ethics* (Cambridge, 2014), p. 291.
- 21 EN 1175b16-18
- 22 EN X.7 1177a23-25
- 23 EN X.7 1177a40-41
- 24 EN X.4 1175a

- 25 EN X.4 1175a4-5
- 26 EN X.7 1177b28-30
- 27 EN X.1174b36-37

## Eternalism and Presentism

Jacob Beck

**Abstract:** In this paper I will argue that past and future times do not exist by arguing for Presentism over Eternalism. First, I will answer two common objections to Presentism, known as the truthmaker objection and the objection from special relativity. Then, using my answer to the truthmaker objection, and establishing a dilemma, I will argue that Presentism and Eternalism agree on the set of objects that truth depends upon; they simply refer to the set by different names. Eternalism and Presentism agree about what presently exists and agree about what doesn't presently exist. The Eternalist simply calls the state of these non-presently-existing things existence simpliciter, whereas the Presentism refers to this state as a lack of existence. But Presentism succeeds in referring to the class of objects in a better way. Because of this, I will show that special relativity, if it is a problem for Presentism, is also a problem for Eternalism, and to resolve this problem, we should accept neo-Lorentzian theory. Together, this makes a more coherent and intuitive description of reality without ad hoc explanation. It does this largely by maximizing virtues such as explanatory power and simplicity.

Eternalism is the claim that the past, present, and future all exist. Presentism is the claim that the only time that exists is the present. From this it follows that the only things that exist, exist right now. In other words, the only things that exist are presently existing things. In this paper I will argue that past and future times do not exist by arguing for Presentism. First, I will answer two common objections to Presentism, known as the truthmaker objection and the objection from special relativity. Then, using my answers to these objections as a starting point, I want to give some motivation for the tensed theory of time (since Presentism implies it) and Presentism itself, by claiming this picture (i.e. the denial of truthmaker and the objection from special relativity, the tensed theory of time, and Presentism) makes a more coherent and intuitive description of reality. Furthermore, it does this largely by maximizing virtues such as explanatory power and simplicity.

First, the truthmaker objection goes as follows:

- (1) truth supervenes on what exists,
- (2) there are true facts about the past,
- (3) the truth of these facts does not supervene on what presently exists, therefore
- (Conclusion) what exists must be more than what presently exists.

What premise (1) means is that if there is another world that is identical to ours in terms of what exists,

then all propositions in it have the same truth values. The motivation for premise (1) is that truth depends on the contents of reality. This seems, at least, initially plausible. The motivation for premise (2) is intuitive and seems to be something no one wants to give up. If premise (2) goes, then we cannot say such things as “dinosaurs existed” is true. The motivation for premise (3) is a bit more confusing, so I will focus on that in the next paragraph. As for the conclusion, it does directly refute the main claim of Presentism, and it seems to me to be necessary given the premises.

The motivation for premise (3) is the suggestion of an elaborate hoax world and the truncated past world. Take any fact about the past. Let’s say that the Roman Empire existed. Assume that this is true. Now imagine a world in which this is not true, but all of the same evidence for the Roman Empire exists. Some organization of history-hating aliens planted all the evidence we have of the Roman Empire – weapons, statues, records, etc. How do we know we are not in such a world? Is it not possible? It seems that what presently exists in this hoax world and in our world are exactly the same, yet they differ in terms of facts regarding the past – namely that the Roman Empire existed. Therefore, it seems that what presently exists in a world does not make facts about the past true – i.e. facts about the past do not supervene on what presently exists. There is a problem with this suggestion, however.

The problem with the elaborate hoax, some may argue, is that it is not possible for everything about our world to be the same in this hoax. Some things – perhaps the subatomic structure of the statues – will be off. What presently exists in the two worlds is not the same. The Eternalist can respond to this suggestion, however, with the truncated past world. Imagine a world that popped into existence one second ago. It is exactly the same as our world in terms of everything it contains. Now, it seems that what presently exists in this world and our world are exactly the same. Yet, in this truncated past world, there never was a Roman Empire. In fact, none of the things that are true of the past in our world are true of the past in this world. So it cannot be that truth supervenes on what presently exists.

This motivation seems to be flawless to me. But I still think it appropriate to reject the truthmaker argument. I will do this by attacking premise (1). I will show an equally legitimate way premise (1) could be false, and, because the argument provides no motivation for the truth of premise (1), I will show that the argument fails. I will do this by denying that truth supervenes on what exists, because I will insist that truth does not depend only on what exists, but also on what existed. That is, truth about the past depends on what used to exist. (For example, it is true that the Roman Empire existed because it was true that it exists.)

This answer at first may seem similar to positing the existence of Lucretian properties – i.e. properties of present things about how they were. However, it is not. Positing Lucretian properties attacks premise (3); it posits presently existing things for truth to be based off of. But this adds some unnecessary entities into our metaphysics. It says that facts about what used to exist in the past depend on this extra existent property of our world about how it used to be. But why say that when you can just say that facts about what used to exist in the past depend on what used to exist? Of course what used to exist is a fact that concerns the world, but we don't need to make up new metaphysical properties this way.

One may object that it does not make sense to talk about something that does not exist. But, in answer, it should be noted that according to the Eternalist, truth supervenes on things that do not presently exist, but on things that used to presently exist. This should sound just as bad as saying that under the Presentism theory, truth supervenes on what used to exist. In fact, they are referring to the same set of objects, but by different names! They agree upon what objects truth supervenes, and they pick out the same set. The only thing Eternalism adds is saying that these past things exist simpliciter. But just saying something exists when it isn't physically embodied or located somewhere and cannot physically interact with reality doesn't add any metaphysical entity for truth to depend

on. And we don't need one! There it seems there is an equally legitimate way premise (1) could be false and the truthmaker objection fails.

Now, starting by considering my response to the truthmaker objection that Presentism and Eternalism are saying the same thing about the relation between truth and existence, I want to argue that Presentism is better. I want to argue that even though they both ultimately say the same thing about the relation between truth and objects, Presentism says it better. I will now create a dilemma where the end result is the eternalist must be inconsistent, or say the same thing as Presentism less simply.

Assume for a second that the Eternalist is committed to the following:

(a) Dinosaurs do not presently exist

This will be the first option in our dilemma. Now we can set up an inner dilemma (in the greater dilemma) for the Eternalist. Suppose he says this means that "Dinosaurs do not exist in the present." Then this entails the weaker claim, "Dinosaurs do not exist if it is the present." But then we can establish by the true premise, "it is the present," and modus ponens that "Dinosaurs do not exist." Thus we can establish Presentism. Now suppose that he goes with his other option in this inner dilemma and says, "Never mind; dinosaurs do exist in the present." Now he runs into a problem here too.

He has said that Dinosaurs do presently exist, but Dinosaurs do not exist in the present. This implies that “exist in the present” is not the same as “presently exist,” but they are clearly the same. So it seems that so long as the Eternalist is committed to (a), he runs into a dilemma where he must either accept Presentism (and contradict himself) or say something false. Thus this first horn of the greater dilemma leads to inconsistency.

Now let’s assume that the Eternalist is not committed to (a). (This is the second option in the greater dilemma). Instead he decides he needs to be more specific and he commits himself to:

- (a1) Dinosaurs do not presently presently-exist
- (a2) Dinosaurs do presently exist-simpliciter

By drawing this distinction, the Eternalist has found a way out. Or so it would seem.

If the Eternalist draws the distinctions he must (because the first horn of the dilemma is inconstant, which is unacceptable) and says that Dinosaurs do exist-simpliciter in the present, he is internally consistent, but he has said too much. Or rather, he has confused matter by saying things in a way he ought not to. What we started out wanting to say is whether Dinosaurs presently exist. And we wanted to say, “They no longer presently exist!”. But he cannot say this. He must say things like, “They no longer presently presently-exist. Sure we could do this and be internally consistent,

but do we really want to? Do we want to make this distinction every time we say “exist?”. I think the answer is no. We can say something exists simpliciter if it used to exist, but it seems that this does not capture what we mean by the term “exist.” We could make up a new sense of “exist” – i.e. exist across all time – but why should we? It is only the Presentist that does not have to.

In still other words, the Eternalist cannot say, “all times exist simpliciter, but (currently) only the present exists,” because “exists” includes “exist simpliciter.” He needs to say something like, “all times exist-simpliciter, but only the present presently-exists,” but this is awkward and unnecessary. Thus, Presentism and Eternalism agree on the set of objects that truth supervenes upon; they simply refer to the set by different names. Eternalism and Presentism agree about what presently exists and agree about what doesn’t presently exist. The Eternalist simply calls the state of these non-presently-existing things existence simpliciter, whereas the Presentism refers to this state as a lack of existence. But Presentism succeeds in referring to the class of objects in a better way.

Next, I want to talk about the objection from special relativity. Picture a moving train with a light bulb placed in the center of one of the carts. The light is turned on and light from the bulb heads both toward the front and back of the cart. To an observer in the cart, the front and back walls of the cart are stationary.

Now, according to special relativity, light always travels at a constant speed, no matter how fast the source that emitted it is traveling (which is the bulb in this case). Thus, to an observer in the cart, the light hits the front and back walls simultaneously. From where an observer outside of the train is standing (i.e. his reference frame), however, the rear wall of the cart rushes up to meet the light, and the front wall rushes away. Thus, the light hits the rear wall first, then the front wall. Thus, according to special relativity, there is no fact of the matter about what events occur simultaneously (or absolute simultaneity), but only simultaneity relative to a reference frame. Now, the objection follows: Presentism says that what exists is what exists presently, but what events are in the present relies on them all being simultaneous, and simultaneity depends on the reference frame, and so there cannot be a fact of the matter about whether an event exists and what is present. (E.g. for someone travelling really fast, they will pick out a different set of events as being simultaneously present).

Now, there are two ways people typically reply to this. The first is by suggesting a different way to pick out things as being present in some way other than with some sense of simultaneity (such as all events happening at the same spacetime point or in a light cone). This, however, changes our conception of the present entirely and completely undermines our motivation for (and conception of) Presentism, so I will

not discuss them for the same reason I will not discuss viewing Presentism as the view that the only thing that exists are future dinosaurs. Additionally, the choice for definitions and the location of such points seems arbitrary. So, I will take the second route: insisting that there is a fact of the matter about what happens at the time. This way, we can still have a coherent picture of the present. I will do so by embracing neo-lorentzian theory. According to this theory, there is a privileged reference frame that determines facts about absolute simultaneity, even if we cannot measure it. This theory makes the same predictions special relativity does, but also allows us to maintain our notion of the present.

One may wonder which reference frame, then, is correct. Well, the reference frame of the universe – the stationary one, the one in which nothing is moving. This seems like it may not help, since we cannot find it, but this should be no more of an issue for neo-lorentzian theory than it was for absolute space. And, in fact, this does require absolute space. One may object also that neo-lorentzian theory is an ad-hoc fix for Presentism, but I will show quite the opposite: assuming absolute space does not exist is the ad-hoc fix.

Some may argue that neo-lorentzian theory is an ad-hoc fix, but it actually fits well with absolute space, which does not require an ad-hoc fix as assuming relative space does. Here's why: Imagine a bucket spinning in space with water. The bucket spins, then the water does and forms a concave shape. Now

remove all other matter in the universe. The water is still concave, but is not spinning relative to the bucket, or anything else discernable. So why is the water concave? It is spinning with respect to absolute space. This is Newton's Bucket argument. Mach did provide an answer to this argument – namely that the water would not form this concave shape; this shape only occurs when the bucket is rotating with respect to other matter (most of the mass in the universe in fact). But to me, it doesn't seem right that something outside of the bucket can effect the accretion it feels. Also, it seems wrong that the only difference between a bucket spinning and not spinning could be the presence of a large stationary rock next to it. Moreover, that this rock soaring around the bucket would also make the bucket feel acceleration and the water change shape. To posit otherwise, as Mach does, seems like an ad hoc fix to the laws of nature to me. Thus Presentism forms a nice coherent picture with neo-L theory and absolute space. Furthermore, Eternalism actually faces the same issue with special relativity.

Recall what we learned from the truthmaker objection – namely, that Presentism and Eternalism pick out the same class of objects as existing and presently-existing respectively. This brings up an interesting point about special relativity. Since Presentism and Eternalism do agree on what presently-exists, then isn't special relativity just as much a problem for Eternalism? Special relativity would imply for Eternalism that what

presently-exist does not depend on any one present; it is also relative to a reference frame. Thus the same argument can be made: Eternalism says that what presently-exists is what presently-exists presently, but what events are in the present relies on them all being simultaneous, and simultaneity depends on the reference frame, and so there cannot be a fact of the matter about whether an event presently-exists. (E.g. for someone travelling really fast, they will pick out a different set of events as being simultaneously present. In one reference frame, things in our past could presently-exist). There is still an issue here to be resolved. This provides more reason to think that the neo-lorentzian theory is the correct one, and that special relativity is not an issue for Presentism.

Additional motivation for the picture I am presenting can be seen in the tensed theory of time, which Presentism implies. One may worry that this complication is a problem for Presentism, but I hold it to be, in fact, an explanatory virtue. When a theory is underdetermined by empirical evidence (i.e. there are multiple theories consistent with the data), it is commonly held that we have the epistemic right to choose by maximizing the supraempirical virtues, and one of these is explanatory power. First, let me discuss the tenseless and tensed theories of time.

The tenseless theory of time is the claim that maintains there are no irreducibly tensed facts – i.e. there is no fact about what is happening right NOW in

an objective sense, nor about what time is the present. According to this theory, when we say “now,” we mean something more like, “simultaneous with this utterance” or “simultaneous with this thought.” The tensed theory is the negation of the tenseless theory of time – i.e. there is a fact about what is happening right NOW.

The first motivation for the tensed theory I want to give is based on the intuition that time is not static. To me, at least, it seems intuitive that a static timeline depicting the world without a moving present is incomplete and wrong. What would time be if it didn't have a moving objective NOW? How could it flow? Why do we only experience the present? If time doesn't pass, then what even is time? Of course, the Eternalist could answer these questions in some way and even suggest that there is an objective NOW (i.e. combining Eternalism with the tensed theory of time), but in answering them or positing this combination of theories seems like an ad hoc fix. (This combination is known as the moving spotlight view.) Only Presentism is strengthened by the connection to the tensed theory of time. This happens because it explains it. There is an objective NOW because it is the only time that exists.

The second motivation for tensed theory I want to give is from temporal value asymmetry. The temporal value asymmetry refers to the preference people have for painful things to be in the past and pleasant things to be in the future. For example, imagine on Monday

you know that you have a painful doctor's the next day – Tuesday. You are deathly afraid of needles greater than 0.2 inches and this one is going to be a whopping 0.5 inch needle. You don't want this appointment to ever come – you don't want it to be in your future – but you go through with the necessary appointment, terrified. Afterward, the pain and fear are gone. You don't care about the appointment anymore and don't dread it's passing. You prefer for it to be in the past, and not the future. Additionally, the same thought process can be mirrored for future events. Taken together, we now have rational reasons for our preferences known as the temporal value asymmetry.

Now the tensed theory explains this temporal value asymmetry well: What makes this true is that there is a fact about what WILL happen, a fact about what is happening NOW, and a fact about what HAS happened. Without these facts, we would have nothing to like or dislike. This, however, is precisely what the tenseless theory claims to be false. The tenseless theory claims that these facts are not true. Again, the Eternalist could object and explain temporal value asymmetry by positing the tensed theory, but this is an ad hoc fix. And even the tenseless theorist could explain the temporal value asymmetry by saying it is irrational, but then it is on the proponent to posit another cause for the non-arbitrary pattern of our preferences, and, assuming one exists, the theory will become more complex (and less simple).

So, now we are left the fact that Presentism explains the tensed theory of time, while Eternalism does not. Additionally, the tensed theory of time explains the temporal value asymmetry, so Presentism does as well, in turn. Thus, due to the tensed theory of time, it seems, the explanatory power of Presentism is greater than that of Eternalism.

There is, additionally, one more supra empirical virtue that Presentism maximizes, that I have hinted at. This is simplicity. The argument here is fairly straightforward: Presentism only posits the existence of one thing – the Present – while Eternalism also requires the existence of the past and future. Since Presentism maximizes this virtue, and is preferable, or at least equal to Eternalism in others (such as predictive power), Presentism is preferable and the theory that ought to be accepted.

In conclusion, in this paper I argued that past and future times do not exist by arguing for Presentism. I answered two common objections to Presentism – truthmaker objection and the objection from special relativity. Then, using my answer to the truthmaker objection as a starting point, I argued that Presentism and Eternalism don't actually disagree about what truth depends on, and that Presentism says it better. Then assuming this, I showed that it elucidates an issue for Eternalism under special relativity and why we should accept neo-lorentzian theory. Additionally, I showed that that in light of this picture which fits well

together without ad hoc explanation (such as assuming absolute space does not exist), it seems Presentism is the correct theory. And, not only that, but Presentism also maximizes supra empirical virtues of explanatory power and simplicity in its explanation of the tensed theory (and in turn the temporal value asymmetry) and in its lack of unnecessary additions, namely the existence of the past and future. Thus, Presentism explains things we have motivation to believe and is simpler. The Present, it seems, is the only thing that exists – a description of what exists and is happening NOW.

