

**volume VIII ♦ issue 1**  
**2017**

# Logos & Episteme

an international journal  
of epistemology

**Romanian Academy**  
**Iasi Branch**



**“Gheorghe Zane” Institute**  
**for Economic and Social**  
**Research**

## Editorial board

---

### Director

Teodor Dima

### Editor-in-Chief

Eugen Huzum

### Executive Editors

Vasile Pleșca

Cătălina-Daniela Răducu

### Assistant Editors

Irina Frasin

Bogdan Ștefanachi

Ioan Alexandru Tofan

### Web&Graphics

Codrin Dinu Vasiliu

Virgil-Constantin Fătu

Simona-Roxana Ulman

---

### Contact address:

Institutul de Cercetări

Economice și Sociale „Gh.Zane”

Iași, str.T.Codrescu, nr.2, cod 700481

Tel/Fax: 004 0332 408922

Email: logosandepisteme@yahoo.com

[www.logos-and-episteme.proiectsbc.ro](http://www.logos-and-episteme.proiectsbc.ro)

## Advisory board

---

Frederick R. Adams

University of Delaware, USA

Scott F. Aikin

Vanderbilt University, USA

Daniel Andler

Université Paris-Sorbonne, Paris IV, France

Alexandru Boboc

Academia Română, Romania

Panayot Butchvarov

University of Iowa, USA

Mircea Dumitru

Universitatea din București, Romania

Sanford Goldberg

Northwestern University, Evanston, USA

Alvin I. Goldman

Rutgers, The State University of New Jersey, USA

Susan Haack

University of Miami, USA

Stephen Hetherington

The University of New South Wales, Sydney, Australia

Paul Humphreys

University of Virginia, USA

Jonathan L. Kvanvig

Baylor University, USA

Thierry Martin

Université de Franche-Comté, Besançon, France

Jürgen Mittelstraß

Universität Konstanz, Germany

Christian Möckel

Humboldt-Universität zu Berlin, Germany

Maryvonne Perrot

Université de Bourgogne, Dijon, France

Olga Maria Pombo-Martins

Universidade de Lisboa, Portugal

Duncan Pritchard

University of Edinburgh, United Kingdom

Nicholas Rescher

University of Pittsburgh, USA

Rahman Shahid

Université Lille 3, France

Ernest Sosa

Rutgers University, USA

Alexandru Surdu

Academia Română, Romania

John F. Symons

University of Texas at El Paso, USA

# TABLE OF CONTENTS

## RESEARCH ARTICLES

Guy AXTELL, Thinking Twice about Virtue and Vice: Philosophical Situationism and the Vicious Minds Hypothesis.....	7
Umut BAYSAN, A New Response to the New Evil Demon Problem.....	41
John BIRO, Non-Pickwickian Belief and ‘the Gettier Problem’.....	47
Olga RAMÍREZ CALLE, Tracing the Territory: A Unitary Foundationalist Account.....	71

## DISCUSSION NOTES/DEBATE

Murray CLARKE, Fred ADAMS, and John A. BARKER, Methods Matter: Beating the Backward Clock.....	99
Jonathan L. KVANVIG, Reply to Simion.....	113
Kevin McCAIN, Undaunted Explanationism.....	117
Moti MIZRAHI, Why Gettier Cases Are Still Misleading: A Reply to Atkins.....	129
Asbjørn STEGLICH-PETERSEN, Weighing the Aim of Belief Again.....	141
Ema SULLIVAN-BISSETT and Paul NOORDHOF, Another Defence of Owen’s Exclusivity Objection to Beliefs Having Aims.....	147
Thomas WILK, Inferences, Experiences, and the Myth of the Given: A Reply to Champagne.....	155
Notes on the Contributors.....	163
<i>Logos and Episteme</i> . Aims and Scope.....	167
Notes to Contributors.....	169



## **RESEARCH ARTICLES**



# THINKING TWICE ABOUT VIRTUE AND VICE: PHILOSOPHICAL SITUATIONISM AND THE VICIOUS MINDS HYPOTHESIS

Guy AXTELL

**ABSTRACT:** This paper provides an empirical defense of credit theories of knowing against Mark Alfano's challenges to them based on his theses of *inferential cognitive situationism* and of *epistemic situationism*. In order to support the claim that credit theories can treat many cases of cognitive success through heuristic cognitive strategies as credit-conferring, the paper develops the compatibility between virtue epistemologies *qua* credit theories, and dual-process theories in cognitive psychology. It also provides a response to Lauren Olin and John Doris' 'vicious minds' thesis, and their 'tradeoff problem' for virtue theories. A genuine convergence between virtue epistemology and dual-process theory is called for, while acknowledging that this effort may demand new and more empirically well-informed projects on both sides of the division between Conservative virtue epistemology (including the credit theory of knowing) and Autonomous virtue epistemology (including projects for providing guidance to epistemic agents).

**KEYWORDS:** bounded rationality, dual-process theory, ecological rationality, heuristic reasoning, situationism, virtue epistemology

## 1. The Great Rationality Debate

One goal of this paper is to defend virtue epistemology (hereafter VE) against a number of charges that Mark Alfano brings against it based upon the incompatibility of its claims with epistemic situationism, and that Lauren Olin and John Doris bring against it based upon an objection we can call the 'trade-off' problem.<sup>1</sup> The latter problem alleges a dilemma in the form of a necessary tradeoff with ill-consequences for VE, a trade-off between the normative appeal of a virtue-theoretic (ability) condition on knowing, and the empirical adequacy of such a condition. Doris presented a short version of the trade-off problem in his

---

<sup>1</sup> Mark Alfano, *Character as Moral Fiction* (Cambridge: Cambridge University Press, 2013); Lauren Olin and John M. Doris, "Vicious Minds," *Philosophical Studies* 168, 3 (2014): 665-692; John Doris, *Lack of Character* (Cambridge: Cambridge University Press, 2002).

critique of virtue ethics in *Lack of Character*, Olin and Doris develop the problem, extending it into a challenge for virtue epistemologies as well.

At the same time, however, the paper will have more constructive goals than that of defending VE against the challenge of philosophical situationists and their close cousins. Epistemologists have not paid enough attention to what psychologists call the *normative-descriptive gap*, or to the ‘bounded’ or ‘ecological’ nature of human cognition. Our human susceptibilities to motivational and cognitive biases may be well-recognized by today’s more naturalistically-inclined philosophers. But there remains a worry about normative theories in ethics and epistemology, shared by many in the social scientific community, which must be a concern for virtue theories as well: philosophers must be warned of repeating the errors of the past by permitting discussions of ethical and epistemic normativity to continue as a ‘separate culture’ from the social and cognitive sciences. As Appiah puts the point, “The questions we put to the social scientists and physiologists are not normative questions. But their answers are not therefore irrelevant to normative questions.”<sup>2</sup>

Naturalistic virtue theory agrees: It is a dubious non-naturalism in philosophy, and in particular an intellectualist tradition associated with the autonomy of philosophy from the human and natural sciences, that motivates the separate cultures notion. Virtue theories have arguably been at the forefront of the movement to integrate the scientific image of humans into normative ethics and epistemology. But my approach suggests that much more attention needs to be paid by ethicists and epistemologists both to the normative-descriptive gap and more particularly to how to conceive the relationships between ‘reasoning and thinking,’ ‘competence and performance,’ and ‘assessment and guidance.’ In terms of positions in the great debate over human rationality which has raged over the past half-century, I will try to show virtue theorists as Meliorists standing against the situationist’s Skepticism. But avoiding the traps of intellectualism and the philosophy-as-autonomy view also means distancing one’s stance from those proponents of the Apologist and Panglossian positions who Daniel Kahneman chastised decades ago as acknowledging only two categories of errors, “pardonable errors by subjects and unpardonable ones by psychologists.”<sup>3</sup>

---

<sup>2</sup> Kwame Anthony Appiah, *Experiments in Ethics* (Cambridge: Harvard University Press, 2008), 62.

<sup>3</sup> Daniel Kahneman quoted from Keith Stanovich and Richard West, “Individual Differences in Reasoning: Implications for the Rationality Debate?” in *Heuristics and Biases: The Psychology of Intuitive Judgment*, eds. Thomas Gilovich, Dale W. Griffin, and Daniel Kahneman (Cambridge: Cambridge University Press, 2002), 421.



Let us be somewhat clearer about what will be the critical and constructive goals of the paper. On the critical side I take the approach of empirical rebuttal and argue that Alfano's two key theses, which he terms *inferential cognitive situationist* (ICS) and *epistemic situationist* (ES) (sections 2 and 3, below) are *generalizations* that are not strongly supported by the selection of studies in cognitive and social psychology (respectively) he bases them on. Generalization as a form of inductive arguments is assessed as either strong or weak, and weak inductive arguments are akin to invalid deductive arguments in falling short of logical compulsion. If such generalizations as philosophical situationism depends upon are not as strong or cogent as they claim, then of course they do not provide a benchmark (as their challenges assume) against which the empirical adequacy of an epistemology should be measured. Sections 2 and 3 respond to challenges directed against the more reliabilist and the more responsibilist versions of VE, respectively. But my *weak generalization* claim about the status of Alfano's two key theses regarding human cognitive agency also extends to Olin and Doris' 'vicious mind' thesis. These authors claim to infer from their discussion of select studies to the 'enormous' variability in human cognitive functioning due to our sensitivity to even slight situational variables. Section 4 offers a fuller response to their 'trade-off' problem for virtue theories than I have previously given, allowing me a chance to more fully develop the *Narrow-Broad Spectrum of Agency-Ascriptions* I alluded to in an earlier exchange.<sup>4</sup>

On the constructive side, I concede that few extant versions of VE have tried explicitly to square themselves with bounded rationality, or with dual-process theory (hereafter DPT), as I will argue that they should. Indeed, there has been more work on accommodating DPT among ethicists than among epistemologists.<sup>5</sup> Bounded rationality, extending from seminal work by Herbert Simon, asks and studies how real people make decisions with limited time, information, and computation. Gerd Gigerenzer writes that the science of

---

<sup>4</sup> Guy Axtell, "Agency Ascriptions in Ethics and Epistemology," in *Virtue and Vice, Moral and Epistemic*, ed. Heather Battaly (Oxford: Wiley/Broadview Press), 73-94. See also Christopher Lepock, "Unifying the Intellectual Virtues," *Philosophy and Phenomenological Inquiry* 83, 1 (2011): 106-128.

<sup>5</sup> Daniel Lapsley and Patrick Hill, "On Dual Processing and Heuristic Approaches to Moral Cognition," *Journal of Moral Education* 37, 3 (2008): 313-332; Holly M. Smith, "Dual-Process Theory and Moral Responsibility," in *The Nature of Moral Responsibility: New Essays*, eds. Randolph Clarke, Michael McKenna, and Angela M. Smith (Oxford: Oxford University Press, 2015), 176-208; Nancy Snow, *Virtue as Social Intelligence: An Empirically Grounded Theory* (London: Routledge, 2009); Snow, "Habitual Virtuous Actions and Automaticity," *Ethical Theory and Moral Practice* 9, 5 (2006): 545-561.

heuristics “Asserts that ecologically valid decisions often do not require exhaustive analysis of all causal variables or an analysis of all possible actions – and – consequences. The best decisions do not always result from such effortful, reflective, calculations, but instead rely on ‘frugal,’ incomplete and truncated inquiry.”<sup>6</sup> Ecological rationality challenges expectations that human reasoners are rational or justified only when they meet normative standards derived independently of empirical and social psychology. It suggests that demands upon rationality be perfectly feasible for agents, computationally speaking, and that norms of epistemic assessment, while still truth connected, not be ‘free-floating’ impositions.<sup>7</sup> That cognition is so heavily ecological means that norms of epistemic rationality and responsibility bump up against pragmatic constraints and inborn limitations in ways that challenge ideal observer and maximizing conceptions of reasoning.<sup>8</sup>

Gigerenzer’s approach has substantial differences from those versions of ecological rationality that I want to highlight in this paper, versions that fall under the umbrella term of dual process theories. No less a pioneer of the biases and heuristics studies than Daniel Kahneman notes that “Tversky and I always thought of the heuristics and biases approach as a two-process theory.”<sup>9</sup> What he describes in his recent book *Thinking, Fast and Slow* as ‘Type 1’ and ‘Type 2’ thinking (hereafter T1/T2) merely modifies terms he acknowledges were introduced first by Keith Stanovich, Richard West and Jonathan St. B.T. Evans, whose work we will focus on in the next section.

---

<sup>6</sup> Gerd Gigerenzer, *Ecological Rationality: Intelligence in the World* (Oxford: Oxford University Press, 2012), 20.

<sup>7</sup> See Rysiew’s helpful conditions on psychology-sensitive philosophical norms of rationality, in Patrick Rysiew, “Rationality Disputes – Psychology and Epistemology,” *Philosophy Compass* 3, 6 (2008): 1153-1176.

<sup>8</sup> Gigerenzer sees the facts of ecological rationality challenging what he terms the ‘classical conception of rationality,’ a conception with ‘appealing but often unrealistic goals’ that he thinks is anti-naturalistic in its tenor yet remains still deeply-rooted in philosophy, economics, and decision theory. The standard view in the cognitive sciences associated with unbounded rationality Gigerenzer blames for the institutionalized division of labor between principles based upon the ‘is’ and ‘ought’ division. “Until recently, the study of cognitive heuristics has been seen as a solely descriptive enterprise, explaining how people actually make decisions. The study of logic and probability, by contrast, has been seen as answering the normative question of how one should make decisions.” (Gigerenzer, *Ecological Rationality*, 496). This split or schism Gigerenzer thinks has served to wrongly elevate logic and probability above heuristics; the result is “contrasting the pure and rational way people *should* reason with the dirty and irrational way people in fact *do* reason.”

<sup>9</sup> Daniel Kahneman, *Thinking, Fast and Slow* (New York: Farrar, Straus & Giroux, 2012).

Since few attempts have already been made to square virtue theory with ecological rationality or with dual-process theory, this paper is one with prescriptive implications for the direction of epistemological research. My title should be taken to reflect the idea that virtue theorists both in ethics and epistemology *do* indeed need to think twice if their appeal to person-level abilities and to characterological concepts more generally is to be empirically-informed and naturalistically sound. Adapting Evans' book title *Thinking Twice: Two Minds in One Brain* (2010),<sup>10</sup> the real and direct challenge to virtue epistemologists as I conceive it, is to rethink the *epistemic credit*-related family of concepts in light of the distinction between (T1) fast, automatic, holistic, and intuitive ways of thinking that require relatively little cognitive effort, and (T2) slow, deliberative, and serial or analytic ones that require substantially more sustained cognitive effort [see **Table 1** reprinted from Evans and Stanovich].<sup>11</sup> Thinking fast/frugally and slow/effortfully are both adaptive; both are routinely successful, although both can also fail. But far too often an effortful, maximizing conception of rationality has been taken as paradigmatic, with heuristic reasoning viewed only as a source of error and dysrationalia. So my thesis of possible convergences between normative epistemology and DPT prescribes substantial rethinking on the part of epistemologists, including proponents of VE.

**Table 1.** Clusters of Attributes Frequently Associated With Dual-Process and Dual-System Theories of Higher Cognition

Type 1 process (intuitive)	Type 2 process (reflective)
Defining features	
<i>Does not require working memory</i>	<i>Requires working memory</i>
<i>Autonomous</i>	<i>Cognitive decoupling; mental simulation</i>
Typical correlates	
Fast	Slow
High capacity	Capacity limited
Parallel	Serial
Nonconscious	Conscious
Biased responses	Normative responses
Contextualized	Abstract
Automatic	Controlled
Associative	Rule-based
Experience-based decision making	Consequential decision making
Independent of cognitive ability	Correlated with cognitive ability
System 1 (old mind)	System 2 (new mind)
Evolved early	Evolved late
Similar to animal cognition	Distinctively human
Implicit knowledge	Explicit knowledge
Basic emotions	Complex emotions

<sup>10</sup> Jonathan St. B. T. Evans, *Thinking Twice: Two Minds in One Brain* (Oxford: Oxford University Press, 2010).

<sup>11</sup> Reprinted from Jonathan St. B.T. Evans and Keith Stanovich, "Dual Process Theories of Higher Cognition: Advancing the Debate," *Perspectives on Psychological Science* 8, 3 (2013): 223-242, 224.

## 2. Abilities, Heuristic Reasoning, and Epistemic Credit

In his book *Character as Moral Fiction* and papers leading up to it, Mark Alfano poses the following dilemma for reliabilist and mixed virtue epistemologies:

First Horn: If they say that heuristics are not intellectual virtues, skepticism looms: If most people use non-virtuous heuristics, then most people have a large number of unjustified beliefs, which do not count as knowledge.

Second Horn: If, however, they say that heuristics are intellectual virtues, then they need to explain how these dispositions are to be construed as reliable.

The dilemma tries to force a choice between holding absolutely either that heuristic reasoning is or isn't virtuous. Virtue epistemologists do not want to set conditions on knowledge so high that knowledge becomes scarce. But embracing the second horn Alfano thinks is barred by the thesis of *inferential cognitive situationism* (ICS):

(ICS) People acquire and retain most of their inferential beliefs through heuristics rather than intellectual virtues.<sup>12</sup>

Arguably, there is a problem with Alfano's statement of the problem since (ICS), by *presupposing* that heuristics are not virtues (the either/or language of 'rather than') begs the question against someone who wants to 'grab the second horn' of the dilemma. But more charitably, since all of Alfano's and Olin and Doris' challenges focus specifically around the aretaic or ability condition that credit (or achievement) theories<sup>13</sup> place upon knowledge possession, I will construe the demand for clarity about the status of heuristics to be about whether virtue epistemologists can accommodate them in a credit theory of knowing. If knowledge is an achievement creditable in significant measure to an agent's manifestation of ability/virtue, how can inferences that apply heuristic reasoning *instead* of patterns that we recognize as sound and reliable inferential practice, count as virtuous?

The issue Alfano raises is indeed important, but the demand he places upon reliabilist and mixed VE seems to me a false dichotomy, so I will go 'Between the Horns' of his dilemma in reply. For Alfano's dilemma clearly takes 'heuristics' and 'cognitive virtues related to deductive and inductive reasoning' out of context from bounded rationality theories. Although bringing concern with heuristic reasoning to the forefront of epistemology can be applauded as likely to spur

---

<sup>12</sup> Alfano, *Character as Moral Fiction*, 201 and 191.

<sup>13</sup> I here take 'credit theory' as a general type of analysis broad enough to include both Robust and Anti-luck VE.

progress in the field, virtue epistemologists can surely avail themselves of leading theories in cognitive psychology, theories on which it would be absurd a demand a choice between the reliability or unreliability of heuristic strategies. Cognitive and social psychology indeed militates against pitting facts about *thinking* against norms of *reasoning* in the way that Alfano's dilemma pits them. To briefly digress, DPT and Gigerenzer's ecological rationality theory have some sharp differences,<sup>14</sup> but it is fundamental to the broader 'new paradigm' in cognitive psychology which they share that intuitive and reflective thinking can each be highly reliable when well-matched to the agent's problem situation. In Evans' terms, "both Type 1 and Type 2 processing can lead to 'good' or normative answers."<sup>15</sup>

Although logocentrism, and a deductivist bias, still casts its shadow on philosophy's ways of approaching norm governance, most epistemologists, and I think all of those associated with VE, are today concerned with human thinking, not just 'reasoning,' and with inference, not just 'argument.' Concerns with success on cognitive tasks through heuristic strategies and T1 processing should indeed prompt a more minimal account of epistemic credit than an epistemology could offer if it remained locked into understanding reasoning and inference only on the model of *argument*. As Paul Thagard argues, a naturalistic epistemology consistent with the successes of ecologically rational agents needs to maintain that "rationality should be understood as a matter of making effective inferences, not just good arguments."<sup>16</sup> But assumptions that might restrict epistemic credit to the

---

<sup>14</sup> Gigerenzer's ecological rationality theory finds unmotivated the division into two types, systems, or 'minds,' arguing instead that intuitive and deliberate judgments are based on common principles such that a unitary rather than a dual-process account can be given of them. But the overlap of shared lessons from cognitive psychology is nevertheless broad. See Arie W. Kruglanski and Gerd Gigerenzer, "Intuitive and Deliberate Judgments are Based on Common Principles," *Psychological Review* 18, 1 (2011): 97-109.

<sup>15</sup> Jonathan St. B.T. Evans, "Dual-Process Theories of Reasoning: Facts and Fallacies," in *The Oxford Handbook of Thinking and Reasoning*, eds. Keith J. Holyoak and Robert G. Morrison (Oxford: Oxford University Press, 2012), 23. See also Evans, "Questions and Challenges for the New Psychology of Reasoning," *Thinking & Reasoning* 18, 1 (2012): 5-31; David Over, "New Paradigm Psychology of Reasoning," *Thinking & Reasoning* 15 (2009): 431-438.

<sup>16</sup> Paul Thagard, "Critical Thinking and Informal Logic: Neuropsychological Perspectives," *Informal Logic* 31, 3 (2011), 152. Manktelow (*Thinking and Reasoning*) and Over ("The New Paradigm") both discuss how seminal to psychology the *gap problem* and the distinction between reasoning and thinking have been. Compare Thagard, writing as a naturalistic epistemologist critiquing the costs of confusing inference and argument: "If inference were the same as argument, it would have the same serial, linguistic structure. However, there's ample evidence from cognitive psychology and neuroscience the human inference is actually parallel

latter are hardly to be associated with reliabilist VE, where naturalistic orientation runs highest. Epistemic credit associated with knowledge-attributions generally demands only weak cognitive achievements unless we are talking about specifically *reflective* knowledge. If this is correct then it can credit successes that might be easily had in untaxing ways in epistemically friendly environments.<sup>17</sup> One would think that the turn in epistemology from states and standings to epistemic agency and inquiry would underline these points. Evans writes,

From a pragmatic viewpoint, even if people fall prey to certain biases, it does not mean they are irrational [or generally unreliable, or ‘vicious’]. Making mistakes can still be part of a rational, or a reliable, or an intellectually virtuous agent’s repertoire. As some leading dual process psychologists have argued more explicitly, errors of thinking occur because of, rather than in spite of, the nature of our intelligence. In other words, they are an inevitable consequence of the way in which we think and a price to be paid for the extraordinary effectiveness with which we routinely deal with the massive information processing requirements of everyday life.<sup>18</sup>

Another way of putting my point is that having an account of inferential knowledge and recognizing virtues connected with sound deduction and abduction does not commit epistemologists to what Adam Morton, in his recent book *Bounded Thinking: Intellectual Virtues for Limited Agents* calls an N-theory of rationality. N-theories derive their norms *independently* of ecological considerations of time, information, and energy, and so often maintain some computationally demanding conception of rationality. Another aspect of the appeal of the naturalistic turn in epistemology, which virtue theory provides an interpretation of, is that we should not dichotomize (as non-naturalistic theories sometimes have) between norms of epistemic assessment and the aim of providing agents with guidance. While naturalistic approaches in epistemology will still recognize that the normativity of assessment and guidance should be distinguished, questions of psychology cannot be treated today as they sometimes

---

rather than serial, multimodal rather than just language-based, and as much emotional as cognitive” (Thagard, “Critical Thinking,” 152).

<sup>17</sup> Compare Pritchard, who identifies strong cognitive achievements with overcoming a significant obstacle to cognitive success, or with the manifestation of high levels of cognitive skill. Weak cognitive successes by contrast are those where it is very easy to attain the relevant cognitive success. Here “one will meet the rubric for cognitive achievements pretty easily.” (Duncan Pritchard, “Virtue Epistemology and the Epistemology of Education,” *Journal of Philosophy of Education* 47, 2 (2013): 236-247, 240.)

<sup>18</sup> Evans quoted in Robert Sternberg and Talia Ben-Zeev, *Complex Cognition: The Psychology of Human Thought*, (Oxford: Oxford University Press, 2001), 194.

have in the past as merely pragmatic or as only guidance but not assessment-related.<sup>19</sup> By focusing on how our intellectual habits, whether genetically-based or acquired, ground our pretensions to knowledge, understanding, or other epistemic goods, virtue epistemologies it seems to me bypass the worries about N-theories and armchair epistemologies that always demand of agents lots of self-reflection and high effort or ‘load-heavy’ cognitive processing.

It is only when an agent applies a cognitive strategy (such as a heuristic pattern of reasoning) *outside* its known range of reliability, or perhaps where System 2 or metacognitive ‘over-ride’ skills and sensitivities were expected of virtuous agents but were not manifested, that epistemologists are likely to balk at credit for success. In these instances error possibilities are high, and the agent’s employed strategy is said to have a low *ecological validity*. Judgments of these sorts may rightly be used to deny epistemic credit (and hence knowledge or understanding). Of course, the instances where T2 over-ride failures occur in humans are many and not few. Philosophers have not come to grips with what Kahneman simply describes as ‘the quiriness of System 1 and the laziness of System 2.’ There are numerous causes of *dysrationalia* in heuristics and biases task experiments that DPT recognizes, including failure to detect the *need* to override a heuristic response, lack of acquired ‘mindware’ available to carry out override, and inability for ‘sustained decoupling’ that allows hypothetical reasoning and other powerful metacognitive aptitudes to engage.

To summarize thus far, I think there is ample opportunity to go ‘Between the Horns’ of Alfano’s dilemma for reliabilist and mixed virtue epistemologies. No one need deny the valuable functions that N-theories serve;<sup>20</sup> but neither should a credit theory of knowing place the agent under any assumed burden of following strategies that maximize cognitive load. VE can and should agree with DPT that,

Since the fast, automatic, and evolutionarily older system requires little cognitive capacity, everyone has the capacity to deal rationally with many reasoning and decision making problems that were important in the environment in which we evolved. Moreover, since the new, slow, rule-based system can be significantly affected by education, there is reason to hope that better educational strategies

---

<sup>19</sup> Michael Bishop, who clearly holds that bounded rationality should deeply impact our approach to epistemic normativity, finds that the received internalist notion of responsibility is what heuristic reasoning especially challenges. (Michael Bishop, “In Praise of Epistemic Irresponsibility: How Lazy and Ignorant Can You Be?” *Synthese* 122, 1 (2008): 179-208, 179. See also Michael Bishop and J.D. Trout, *Epistemology and the Psychology of Human Judgment* (Oxford: Oxford University Press, 2005).)

<sup>20</sup> See Adam Morton, *Bounded Thinking: Intellectual Virtues for Limited Agents* (Oxford: Oxford University Press, 2013).

will improve people's performance on those problems that the old system was not designed to deal with.<sup>21</sup>

Situationists are skeptics about character-traits, and so also of the claims about individual differences among agents that character-traits help determine. But in order to go beyond arguing for a basic consistency of VE with cognitive psychology's recognized continuum between fast and slow, or intuitive and reflective ways of processing, I offer reflections on three specific guiding themes in the work of Evans and Stanovich that I think point out potential *convergences* between them:

**Individual differences.** According to proponents of DPT there are few continuous individual differences among people with respect to autonomous mind, or Type 1 processing. We are all energy economizers and want to fit strategies to problems ecologically when we can, rather than doing the 'expensive' reasoning of ideal inquirers *qua* unbounded reasoners. But they also insist that "the intelligence of the new mind is quite variable across individuals."<sup>22</sup> DPT both predicts and finds confirmed substantial individual differences in what Stanovich terms 'rational thinking dispositions' and therefore recognizes "individual differences as essential components of heuristics and biases research."<sup>23</sup> The cultivation of general intellectual virtues that help attenuate cognitive biases and inappropriate heuristic responses strongly overlaps with what Stanovich and his colleagues term the cultivation of our *fluid rationality*. Fluid rationality describes a range of available *critical reasoning dispositions*. Research shows that those who do poorly on cognitive task tend to be 'cognitive misers' in the way that they process, while those who do better than average exhibit a more desirable collaboration between their T1 and T2 thinking, and as Stanovich has it [Table 2], between our *crystallized* and *fluid* rationality. Some of the imperfect but powerful aptitudes fluid rationality describes are 'resistance to miserly information processing,' and 'absence of irrelevant context effects in decision making.'

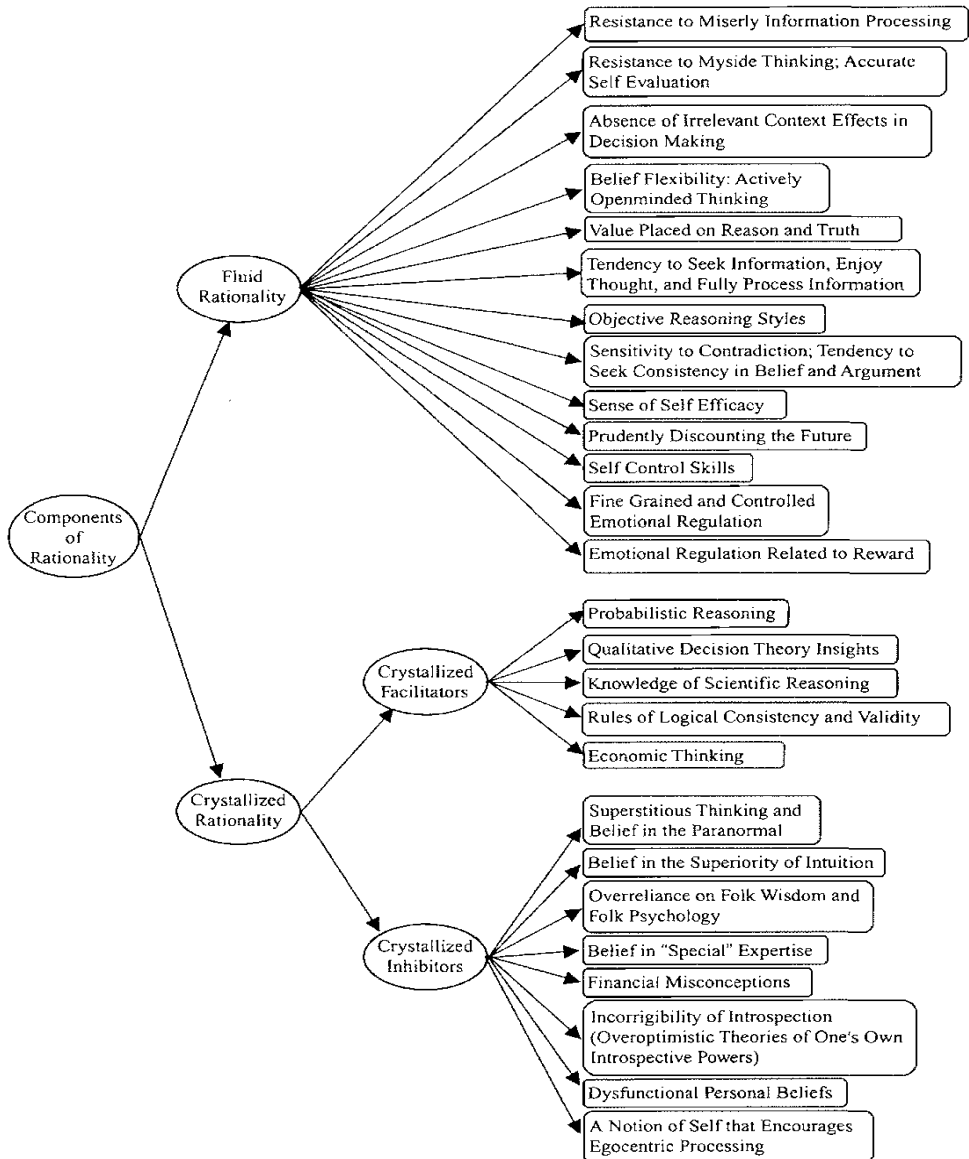
---

<sup>21</sup> Richard Samuels and Stephen Stich, "Rationality and Psychology," in *The Oxford Handbook of Rationality*, eds. Alfred R. Mele and Piers Rawling (Oxford: Oxford University Press, 2004): 279-300.

<sup>22</sup> Evans, *Thinking Twice*, 209.

<sup>23</sup> See also Keith E. Stanovich, Richard F. West, and Maggie Toplak, "Individual Differences as Essential Components of Heuristics and Biases Research," in *The Science of Reason*, eds. Ken Manktelow, David Over, and Shira Elqayam (New York: Psychology Press, 2012), 335-396.



**Table 2. The Conceptual Structure of Rational Thought<sup>24</sup>**

<sup>24</sup> Reprinted from Keith E. Stanovich, Richard F. West, and Maggie Toplak, "Intelligence and Rationality," in *Cambridge Handbook of Intelligence*, eds. Robert Sternberg and Scott Barry Kaufman (Cambridge: Cambridge University Press, 2011), 784-826, 799.

**Critical reasoning dispositions differ from innate IQ and are trainable.** In explaining the significant individual differences in cognitive performance that the heuristics and biases literature evidences, Stanovich urges us to distinguish intelligence from rationality in order to give the proper share to each. For instance, intelligence tests miss much; the magnitude of the myside bias in individuals shows very little relation to intelligence, and avoiding it is one of numerous rational thinking skills that are not assessed by IQ. “[I]ntelligence and rationality occupy different conceptual locations in models of cognition,”<sup>25</sup> and must be measured using different tasks and operations. Individual differences in biases and heuristics tasks are “more related to rationality than intelligence.”<sup>26</sup> Measures of *rational thinking dispositions* are positively correlated with normative performance on tasks and often predict unique outcomes. Of course motivational factors are already in play when we speak of rational thinking *dispositions*. As Evans explains,

For the kinds of problems where reflective reasoning is required, you also need to have the disposition to apply effortful reasoning, rather than to rely on intuitions and feelings. This disposition is partly a matter of personality, but is also influenced by culture and context.<sup>27</sup>

So it is very important from a virtue-theoretic perspective that Stanovich describes many response differences as stemming from differences such as these rather than what IQ tests test for:

[M]ost importantly, IQ is a resource that you have to apply to a problem for it to be of any use to you... so one cause of *dysrationalia* is that while a person of high IQ *could* reason well, they're actually failing to engage their reflective abilities. Instead, they are inclined by their personality or circumstances to rely on gut feelings and intuitions, to be strongly influenced by prior beliefs (which may be false) or prone to social influences by peers (who might not be bright).<sup>28</sup>

So central is this distinction between intelligence and rationality to Stanovich's version of DPT that he quite recently makes a 'tripartite proposal' on which Type 2 reasoning subdivides into the *algorithmic mind* and the *reflective mind*. This subdivision captures the platitude that being intelligent does not guarantee being smart [see Table 3]. If you lack the required critical reasoning

---

<sup>25</sup> Research shows “there is a positive correlation between IQ and rational thinking, but it is relatively modest in size...people vary not only in their cognitive ability but also in their disposition to think critically about problems they face...” (Keith Stanovich, *Rationality and the Reflective Mind* (Oxford: Oxford University Press, 2011), 206).

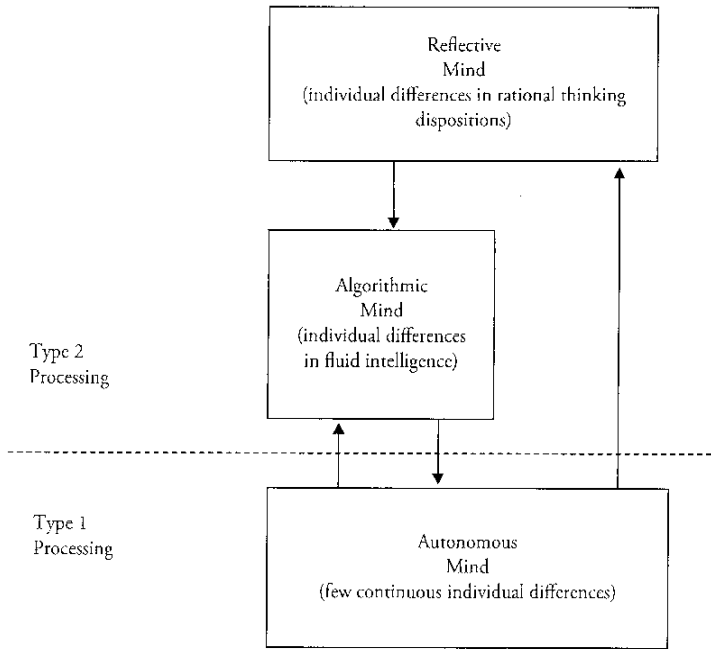
<sup>26</sup> Stanovich, *Rationality and the Reflective Mind*, 206.

<sup>27</sup> Evans, *Thinking Twice*, 210.

<sup>28</sup> Stanovich, *Rationality and the Reflective Mind*, 206.

dispositions or *mindware*, you will find it hard to escape Type 1 intuitive, heuristic responses, even if highly intelligent.

**Table 3. The tripartite structure and the locus of individual differences<sup>29</sup>**



**Meliorism and a balance of inner and outer.** VE and DPT motivate a moderately ‘Meliorist’<sup>30</sup> position in what Stanovich and West call the Great Rationality Debate. Meliorism contrasts with overtly Skeptical automaticity, ‘vicious mind,’ or situationist claims on the one hand, and Apologist/Panglossian views on the other. Stanovich writes that, “What has been ignored in the Great Rationality Debate is individual differences,” something which he cites as devaluing the Meliorist position that proponents of DPT support. Skeptics and Panglossians share an unfounded bent towards underestimating or ignoring the degree of difference found in subject responses in the heuristics and biases studies. Melioristic attitudes

<sup>29</sup> Reprinted from Keith Stanovich, “On the Distinction between Rationality and Intelligence: Implications for Understanding Individual Differences in Reasoning,” in *The Oxford Handbook of Thinking and Reasoning*, 352.

<sup>30</sup> See especially Stanovich, West and Toplak, “Intelligence and Rationality”; and Stanovich and West, “On the Relative Independence of Thinking Biases and Cognitive Ability,” *Journal of Personality and Social Psychology* 94 (2008): 672-695.

of piecemeal improbability of human reasoners are made more plausible by the family of dual processing theories. The point is not unrelated to how Nancy Snow concedes that virtues “might start out by being local,” while holding that “they need not remain so,” and how making knowledge and skills *chronically accessible* through training and habituation is often possible.<sup>31</sup> Kahneman and Frederick for example explore how “complex cognitive operations eventually migrate from System 2 to System 1 as proficiency and skill are acquired.”<sup>32</sup> The automation or chronic accessibility of what was once a slow, effortful process he illustrates through the ability of chess masters to very quickly and without great effort assess the merit of chess moves.

Relatedly, Meliorists are generally moderate in respect to *how* to improve cognitive performance, balancing what might be termed *cognitive change* and *environmental change*. DPT’s prescriptive upshot is that of a *balance* of “teachable reasoning strategies and environmental fixes”; improve the environment where that helps to improve rationality, *and* improve individual skills and competences directly through practice and education.<sup>33</sup> Skeptics about character by contrast criticize character education efforts. In discussing the prescriptive upshot of (ES), Alfano proposes retaining virtue-talk in education primarily as something of a motivational white lie that can enhance short-term cognitive performance primarily by raising the effect of mood. Doris’ conclusion in *Lack of Character* still more pessimistically suggests *turning away* from attempts to develop pedagogy for integrated character to something like enlightened situation-management.<sup>34</sup>

In summary, DPT highlights an empirically well-grounded balance between the inner and the outer. It explains the complex interactions necessary for the successful exercise of different kinds of reasoning, and the trainability of T2 rational thinking dispositions. Behavior is seen as a complex function of the two systems or types reasoning working in cooperation and competition with each other. The importance of T2 skills and dispositions for decision making is evident

---

<sup>31</sup> Snow, *Virtue as Social Intelligence*, 37.

<sup>32</sup> Daniel Kahneman and Shane Frederick, “Representativeness Revisited,” in *Heuristics and Biases*, 49-81, 51.

<sup>33</sup> Stanovich, “Distinction,” 359.

<sup>34</sup> Epistemic and moral paternalism *rather than* character cultivation is appealing to Skeptics. Virtue theorists would reject this but could agree that there needs to be a *balance* between managing situations and acquiring virtue (inner/outer management). Of course, managing situations is partly the individual’s job, as the virtuous are often those that wisely avoid temptations, etc. they know they are susceptible to, rather than those who show outstanding continence or willpower.

from the way they monitor T1 responses and potentially correct for the biases that fast, intuitive processes are most vulnerable to. According to Evans,

We actually know quite a lot, from the experiments conducted by psychologists, about when the reflective mind will intervene in decision-making. This will happen more often when people are given strong instructions for rational thinking, for example to engage in logical reasoning and disregard prior beliefs. There will be less intervention when people are given little time to think about the problem, or are required to carry out another task at the same time that requires their attention.<sup>35</sup>

Stanovich puts much the same point by saying that “mindware gaps most often arise from lack of education and experience.”<sup>36</sup>

I conclude that reliabilist VE is not substantially out of accord with DPT, either in terms of individual differences as anticipated and found in test results on heuristics tasks, or in terms of their explanation. The cautious optimism that character epistemology shares with moderate Meliorist social psychology supports the possibilities of substantially improving education for individual rational thinking dispositions. It therefore finds quite congenial the argument that grounds this possibility empirically in the basic distinction between largely innate IQ and largely acquired intellectual habits and skills. Dual process theory can help philosophers address the question of credit worthiness for success of heuristic strategies and T1 processes. At the same time, DPT allows us to reject Alfano’s dilemma regarding VE and inferential knowledge as based upon the false dichotomy, the forced choice between treating heuristics *en toto* as either virtuous or non-virtuous. Finally, together with the other considerations we have raised, it reveals his first key thesis, (ICS), as a *generalization* based upon that false dichotomy (i.e., a principle fallaciously claiming that people employ heuristics ‘rather than’ manifesting reliabilist virtues). The next section turns attention to studies that Alfano uses from social psychology to support a substantially different challenge, one aimed at responsibilist and inquiry-focused forms of VE.

### 3. Responsibility, Reflective Virtues, and Social Environment

When situationists criticize ‘classical’ (neo-Aristotelian) and ‘inquiry-focused’ (Peircean/Deweyan) virtue responsibilism as empirically inadequate, they naturally enough utilize studies from social psychology.<sup>37</sup> The previous section on

---

<sup>35</sup> Evans, *Thinking Twice*, 205.

<sup>36</sup> Stanovich, “Distinction,” 356.

<sup>37</sup> John Dewey was very concerned not just in formal reasoning, but in how we *think*. His account of habit is highly attentive to the philosophical importance of entrenched aspects of our

Alfano's dilemma for virtue reliabilists was cast in terms of epistemic credit restricted to achievement of epistemic goods through an individual's ability, but Alfano's suggestion is especially pertinent to testimonial knowledge and to discussion of reflective intellectual virtues like intellectual courage, humility, and trust.

One proposal I appreciate in Alfano's *Character as Moral Fiction* is a need to move away from treating individuals as the bearers of virtue, to thinking of them as "a triadic relation among the agent, a social milieu, and an environment."<sup>38</sup> But apart from the attempts by some thinkers to put demanding intellectual motivation conditions on knowledge,<sup>39</sup> the suggestion to conceive reflective virtues in this 'triadic' way could be related to a wealth of recent work among responsibilists on collective and group virtues, and beyond this to strong overlaps one finds among character epistemology, social epistemology, and feminist epistemology today.<sup>40</sup> As inquiry-focused VE does not partake of what Alfano describes as Linda Zagzebski's 'classical responsibilism,' I will not defend her conditions on knowing.<sup>41</sup> We will instead attend to Alfano's more general claim that "the intellectual virtues traditionally countenanced by classical responsibilism

---

cognitive architecture, as well as situational factors within our environment. Habit is the fixed routine of activity which normally predominates, often manifesting in behavior in which consciousness may play only a token role. He states that people often know more with their habits, not with their consciousness; Action may take place with or without an end-in-view, and in the latter case, there is simply settled habit. It is only if a problematic situation arises that habit is disrupted and impulse proves inadequate. At this point, if we have the needed flexibility and metacognitive wherewithal, more effortful thinking intervenes to help resolve the problematic situation, or it does not, and the result is likely to be unsatisfactory. The adjustments are only successful as we have the flexibility of mind to apply a strategy of inquiry well-adapted to the particular situation.

<sup>38</sup> Alfano, *Virtue and Moral Fiction*, 146.

<sup>39</sup> Brendel distinguishes between holding that virtues that function as means of acquiring knowledge are rightly described as dispositions, and the attempt to *define* knowledge in terms of these virtuous dispositions. I agree with her description of Linda Zagzebski's form of virtue responsibilism as leading to 'a counter – intuitive and intellectually over-loaded concept of knowledge.' See Elke Brendel, "The Epistemic Function of Virtuous Dispositions," in *Debating Dispositions: Issues in Metaphysics, Epistemology and Philosophy of Mind*, eds. Gregor Damschen, Robert Schnepf, and Karsten Stüber (Berlin: De Gruyter, 2014). I have elsewhere described inquiry-focused VE as holding an 'overlap' model of ethical and intellectual norms, in contrast to Linda Zagzebski's 'reduction' model. Alfano seems to mean something else by inquiry-focused VE than I do.

<sup>40</sup> Work on civic and collective/group virtues and on theory virtues in science apply aretaic concepts without supposing these to be personal traits.

<sup>41</sup> Linda Zagzebski, *Virtues of the Mind* (Cambridge: Cambridge University Press, 1996).

and inquiry responsible are empirically inadequate.”<sup>42</sup> The support he offers for this empirical inadequacy charge is the force of the thesis he calls *Epistemic Situationism* (ES):

(ES) Most people’s conative intellectual traits are not virtues because they are highly sensitive to seemingly trivial and epistemically irrelevant situational influences.<sup>43</sup>

Alfano’s key theses (ICS) and (ES) are broad generalizations that are purported scientific conclusions from psychological experiments. The main studies that Alfano uses to support his generalization (ES) are the famous Asch Line Task studies of the early 1950’s, and its follow-ups. Staying with the approach of strong empirical rebuttal, I now want to show that Alfano’s use of the Line Task studies depends upon a misleading re-interpretation of them. They do not strongly support the lesson he wants and needs to draw from them. Firstly, Alfano does not mention that the original Asch studies on conformity took place on American subjects exclusively, and during the McCarthy Era when lack of conformity was often identified with lack of patriotism and socially ostracized on that basis. Alfano’s interpretation of the results of these studies is furthermore substantially at odds with the experimenters’ own conclusions. I will now argue that his interpretation is at odds with their own statements both about the *variability of responses* found on the Line Task, and about the *explanations* of these differences in response, where Alfano blatantly ignores Asch’s and Milgram’s explanation involving *cultural* variances.

***Variability of responses:*** Asch did conclude that a majority can influence a minority even in an unambiguous situation in which the correct answer is obvious, confirming (versus M. Sherif) that majority influence is stronger than had been thought. But he saw the studies as clearly showing that people are capable of greater or lesser ‘strength’ in resisting peer pressure. “Among the independent individuals were many who held fast because of staunch confidence in their own judgment.”<sup>44</sup> What Asch and his colleagues found to be ‘the most significant fact’

---

<sup>42</sup> Alfano, *Virtue as Moral Fiction*, 185.

<sup>43</sup> Alfano, *Virtue as Moral Fiction*, 162. See also Alfano’s earlier “Expanding the Situationist Challenge to Responsibilist Virtue Epistemology, *Philosophical Quarterly* 62, 247 (2012): 223-249.

<sup>44</sup> Quotations taken from S. E. Asch, “Opinions and Social Pressure,” *Scientific American* 193, 5 (1955): 31-35. Retrieved July 5, 2016, from <http://www.columbia.edu/cu/psychology/terrace/w1001/readings/asch.pdf>. In a later publication, Asch does again say that “errors increased strikingly” among subjects as majority increased, this being measured against control group. Yet

about the fully 25% best-performers who were consistently independent “was not absence of responsiveness to the majority but a capacity to recover from doubt and to reestablish their equilibrium.” They also found that the degree of independence increases with the deviation of the majority from the truth and with “decreased clarity in a situation.” They held that the study “establishes conclusively that the performances of individual subjects were highly consistent”: those who showed independence from or conformity with a majority over which of three presented lines was longest continued the same pattern of response *over time*. This last conclusion from the data deserves special mention since it seems quite *inconsistent* with the claim of (ES) that most people are highly sensitive to situational variables.<sup>45</sup>

What degree of independence is needed to demonstrate significant individual differences is a matter of much dispute among psychologists, but Alfano’s exclusive emphasis on the degree of conformity relative to a control group not exposed to peer pressure presents a one-sided reading, substantially at odds with Asch’s own conclusions emphasizing very substantial individual differences in motivation or ability to resist peer pressure, and high individual consistency over time. Alfano emphasizes that unanimity among confederates ‘produced striking conformity’<sup>46</sup> in test subjects, while Asch himself emphasizes ‘startling differences’ and strong ‘consistency’ among these same subjects. Neither insight is strictly wrong, but neither one is the whole story.

***Explanation of the data:*** Foremost among questions that Asch says his studies raise is the question of the extent to which these ‘startling differences’ can be attributed to ‘sociological or cultural conditions.’ Asch concluded by suggesting that cultural values are among the primary explainers, and that

the tendency to conformity in our society...raises questions about our ways of education and about the values that guide our conduct...anyone inclined to draw

---

he also concludes that “Individuals responded in fundamentally different ways to the opposition of the majority,” and that “Despite the effect of the majority the preponderance of estimates was, under the present conditions, independent of the majority.” (“Studies of Independence and Conformity: A Minority of One against a Unanimous Majority,” *Psychological Monographs: General and Applied* 70 (1956): 1-70, DOI: 10.1037/h0093718. Retrieved March 24, 2016, from <http://psyc604.stasson.org/Asch1956.pdf> (12).) Note also that if the line task is a ‘rapid response’ task, recent DPT research *predicts* that this ‘taxing’ effect on subjects will result in lowered performance, even independently of the ‘majority effect’ that Alfano tries to say undermines the robustness of intellectual courage as a character trait.

<sup>45</sup> Asch, “Studies,” 20.

<sup>46</sup> Alfano, *Virtue as Moral Fiction*, 183.



too pessimistic conclusions from this report would do well to remind himself that the capacities for independence are not to be underestimated.<sup>47</sup>

Stanley Milgram in 1961 conducted follow-up *comparative* studies in part to see if Asch was right about cultural values deeply impacting response to peer pressure. Comparing French and Norwegian subjects, he found the *hypothesis of cultural variation* to be corroborated by his study. While his conclusions were phrased tentatively (“These findings are by no means conclusive”), he reported that “No matter how the data are examined they point to greater independence among the French than among the Norwegians.” Yet when Alfano does mention Milgram’s “Nationality and Conformity,” he leaves its guiding *hypothesis of cultural variation* totally out of his presentation, instead re-interpreting this study in a self-serving way as just re-confirming a lesson about humans taken *collectively*: that character-traits are routinely swamped by epistemically-irrelevant situational variables like majority effect. So he writes that “As in the original Asch study, unanimity produced striking conformity.” But while Alfano cites situational factors like majority size as salient, Milgram’s own conclusion actually highlighted the salience of cultural influences supported by the finding that “in every one of the five experiments performed in both countries the French showed themselves to be the more resistant to group pressure.”<sup>48</sup> The situationist interpretation of Line Task studies fails to account for the importance of cultural influence on conformity, as tested across time (Americans in the Fifties vs. the Sixties, say), or across cultures (French versus Norwegians). Alfano also omits any mention of the prominent 1996 meta-study of Asch-type line judgment tasks, which actually drew upon 133 studies in 17 countries. The Bond and Smith meta-study found that “levels of conformity had steadily declined since Asch’s studies in the early 1950s”; in respect to just the U.S. studies it found “that the date of study was significantly negatively related to effect size, indicating that there has been a

---

<sup>47</sup> Asch, “Opinions,” 34: “At one extreme, about one quarter of the subjects were completely independent and never agreed with the erroneous judgments of the majority. At the other extreme, some individuals went with the majority nearly all the time.”

<sup>48</sup> “Twelve per cent of the Norwegian students conformed to the group on every one of the 16 critical trials, while only 1 per cent of the French conformed on every occasion. Forty-one per cent of the French students but only 25 per cent of the Norwegians displayed strong independence. And in every one of the five experiments performed in both countries the French showed themselves to be the more resistant to group pressure.” (Stanley Milgram, “Nationality and Conformity, *Scientific American* 205, 6 (1961). Retrieved July 3, 2016 <https://www.scientificamerican.com/article.cfm?id=milgram-nationality-conformity&print=true>.)

decline in the level of conformity.”<sup>49</sup> The section comparing national and national trans-national studies concluded that

the impact of the cultural variables was greater than any other, including those moderator variables such as majority size typically identified as being important factors. Cultural values, it would seem, are significant mediators of response in group pressure experiments.

Here again, what these (non-situationist) authors conclude to be the variable whose impact on task response was ‘greater than any other’ is allowed no significant role in the situationist re-interpretation of these same studies.

These are all reasons to think that the Asch Line Task studies do not provide the evidence needed to make Alfano’s thesis (*ES*) a strong generalization from social psychological experiments. We should for the same reasons of weak empirical support reject Alfano’s associated claims that “virtues identified by inquiry responsibilism are not the sorts of traits that many people possess,” and that “rather than being *intellectually courageous*, people are at best *intellectually courageous unless faced with unanimous dissent of at least three other people*.” It appears that theoretical orientation deeply affects the interpretation that situationists like Alfano apply to the studies they cite and discuss. Add to this that the studies they cite are highly selective and often quite dated ones pulled from a much larger literature on biases and heuristics. Both points greatly diminish the support these studies can provide to the situationist’s generalizations about human behavior.

To digress for a moment, one point that critics have made against behaviorist and situationist methodology is that individual differences in personality or character-related effects routinely get ‘construed as noise’ and therefore deprived of due consideration.<sup>50</sup> Recognition of quite significant

---

<sup>49</sup> Rod Bond and Peter Smith, “Culture and Conformity: A Meta-Analysis of Studies Using Asch’s Line Judgment Task,” *Psychological Bulletin* 119, 1 (1996): 111-137, 124-125, <http://psycnet.apa.org/?&fa=main.doiLanding&doi=10.1037/0033-2909.119.1.111>.

<sup>50</sup> John Kihlstrom argues that understood historically, it is because of what situationism inherits from behaviorism that situationists are predisposed to interpret social psychological studies largely in terms of behavioral reflexes upon a set human nature, or less strongly, in terms of social behavior varying as a function of features of the external environment. Kihlstrom observes that “In such research, the effects of individual differences in personality are generally construed as ‘noise.’ This view is captured by what might be called the *Doctrine of Situationism*: ‘Social behavior varies as a function of features of the external environment, particularly the social situation, that elicit behavior directly, or that communicate social expectations, demands, and incentives’.... Situationism has its obvious origins in stimulus-response behaviorism.” (John

differences in performance, whether attributable more to individual or to cultural character, threatens their conception of human nature as varying largely only as a function of features of the external environment. Alfano for his part calls it a bogeyman to identify situationists with a crude behavioristic model of conduct; he claims that it holds appropriate place for deliberation and sensitivity to reasons in the determination of conduct in an agent. But that Alfano ‘blinks’ over all matters related to the salience of cultural values in the explanation of differences in response among test subjects suggests otherwise. I argue that it clearly shows an inadequate acknowledgment of social milieu, and in this sense betrays their avowed interactionist, triadic relationship between agent, social milieu, and environment. To cross the line between social and cognitive psychology for a moment, we earlier discussed DPT as defending the reality and importance of individual differences (Section 2, above). But the role of cultural values in explaining the variability of responses is easily accommodated and indeed supported in DPT. As Manktelow summarizes in *Thinking and Reasoning*, studies affirm that the new mind is more heavily influenced by culture and formal education. In Stanovich’s terms, people can acquire new ‘mindware,’ and the mindware that people employ is influenced by their culture.

Concern with interpretations of social psychological experiments that ignore the impact of cultural values on task performance highlights the fact that the situationist generalization draws upon a base of mostly latitudinal or ‘single pass’ studies. The lessons that must be learned between psychology and philosophy, this should remind us, run both directions. Methodologies that purport to draw broad generalizations but are based mainly on single-pass heuristic task studies seem to me as open to critique as are studies of moral reasoning narrowly dependent upon intuition-pumping and ‘Trolley-ology,’ or traumatic, dilemma-focused tests. Philosophers and psychologists have both been at fault, and only working together will philosophers and psychologists better address the *Descriptive-Normative Gap problem* introduced at the start of this paper. But before discussing that problem directly in the final section, let me say more about virtue epistemologies themselves, and different forms of they take, since everyone agrees that the situationist challenge affects different forms of VE differently, and some more strongly than others.<sup>51</sup>

---

F. Kihlstrom, “The Person-Situation Interaction,” in *Oxford Handbook of Social Cognition*, ed. Donal E. Carlston (Oxford: Oxford University Press, 2013), 2-3.)

<sup>51</sup> I am taking ‘credit theory’ broadly enough to include both Robust (RVE) and anti-Luck (ALVE) virtue epistemologies. Responsibilist and inquiry-focused VE cut across Jason Baehr’s distinction between autonomous and conservative VE; Inquiry-focused VE may still offer an

Robust, global character-traits and the responsibilist forms of VE that are most concerned with these ‘thick’ characterological and affective concepts are a natural first target of situationist critique. I think of the specific form of VE I support – inquiry-focused VE, or as I have sometimes also termed it, *zetetic* responsibilism in contrast with Zagzebski’s *phronomic* responsibilism – as champion of things *diachronic*. This means it supports and investigates both the backwards-looking concern with the etiology of particular belief, and forward-looking – let us call this *axiological* – concern with the need to reflect upon or improve one’s cognitive strategy and/or evidential situation and to fit one’s epistemic goals, etc.<sup>52</sup> The first is a primary matter for virtue reliabilists and any account of doxastic justification; the second is a concern for both reliabilists and responsibilists, but is not closely connected with epistemic assessment or the project of the analysis of knowledge.

There is a close if not perfect connection between philosophical concern with the diachronic as crucial to both ethical and epistemic normativity, and dissatisfaction with situationist social psychology as generalizing latitudinal studies. There is no dearth of longitudinal studies, but much of it exists over the gap between situationist and automaticity theory, on the one hand, and positive psychology on the other. Blaine Fowers and his colleagues for example try to operationalize Aristotelian virtue theory. It is not incidental to this that in his book *Psychology and Virtue*, Fowers argues that taking latitudinal studies as a basis for ‘judging timelessly’ is highly problematic. If humans don’t just encounter, but also help construct their environments, then situational and agential factors are not as easily sorted out as it might appear when we draw only from latitudinal studies.

If we are serious about exploring whether character strengths actually manifest in behavior, a very different approach to the research [than latitudinal studies take] is necessary. Investigators have to assess three essential components of a trait or character strength. First, there must be individual differences on the

---

account of knowing even while re-envisioning epistemology in the Deweyan fashion of theory of inquiry. For opposing views on whether virtue epistemologies really build more empirical assumptions into their accounts than do other, more generic forms of reliabilist externalism see David Henderson and Terrance Horgan, “Epistemic Virtues and Cognitive Dispositions,” in *Debating Dispositions*, 296-319. Also Christian Nimtz, “Knowledge, Abilities, and ‘Because’-Clauses,” in *Knowledge, Virtue, and Action: Putting Epistemic Virtues to Work*, eds. Tim Henning and David P. Schweikard (London: Routledge, 2013).

<sup>52</sup> See Guy Axtell, “Recovering Responsibility,” *Logos & Episteme* II, 3 (2011): 429-454 for defense of the epistemic centrality of diachronic and not just synchronic evaluations, as internalist evidentialists insist.

characteristic. Second, evidence for consistency in trait-associated behaviors is needed across situations. Third, there must be consistency in trait-associated behavior over time.<sup>53</sup>

Connecting with Fowers second and third point, Jesse Prinz points out that “Factors external to a person can influence behavior in two different ways: synchronically or diachronically,” but that latitudinal studies largely capture only synchronic influence.<sup>54</sup> I find each of these distinctions especially pertinent to the prospects of tying philosophy and the social sciences closer together. For example, sociologist Gabriel Abend argues that if the contemporary science of morality focuses on situations or courses of action judged as ‘right,’ ‘okay,’ or ‘permissible,’ recorded through the push of a button, then it “is not a science of morality, but of thin morality only.” Much as social scientists are rightly skeptical of philosophical ethics presented as dilemma-cases and Trolley-problems, virtue theorists are skeptical of a long tradition in social psychology focused on a ‘thin’ conception of moral and cognitive reasoning.<sup>55</sup>

We can end this section with a passage from Prinz, who argues that the difference between synchronic and diachronic influence matters vitally, because the two forms of influence hold quite different theoretical lessons and implications:

If we were swayed only by synchronic factors, then all people would be the same: put two people in the same situation, and they will probably do the same thing. But, if diachronic influences are possible, then people can internalize social norms, and, as a result, people with different backgrounds will behave differently in the exact same situations. If all people behaved alike in the same situations, character based ethical theories would be in trouble: it would be impossible to *cultivate* character. At best, we could do what Doris recommends: try to put ourselves in situations that promote good behavior. But, if diachronic influence is possible, the cultivation of character *is* possible... Cultivating

---

<sup>53</sup> Blaine Fowers, *Virtue and Psychology: Pursuing Excellence in Ordinary Practices* (Washington DC: American Psychological Association, 2005), 40.

<sup>54</sup> Jesse Prinz, “The Normativity Challenge: Why Empirically Real Traits Won’t Save Virtue Ethics,” *Journal of Ethics* 13 (2009): 117-144, 132.

<sup>55</sup> I have sometimes described my fellow virtue responsibilists as epistemological ‘thickies.’ Abend points out that “Unlike thin predicates, thick predicates have institutional and cultural preconditions or presuppositions ... [T]he moral concepts and properties expressed by those predicates – e.g., the concepts and properties of humanness, gentlemanliness, piousness – are partly constituted by institutional and cultural facts.” (Gabriel Abend, “Thick Concepts and the Moral Brain,” *Archive of European Sociology* LII (2011): 143-172.) For a thin-centered account, see Thomas Hurka, “Virtuous Act, Virtuous Disposition,” *Analysis* 66, 1 (2006): 69-76.

virtuous traits that do not get overwhelmed by synchronic situational variables may be difficult, but there is no reason to think it's impossible.<sup>56</sup>

#### 4. The Normative-Descriptive Gap and the Olin/Doris 'Trade-off' Thesis

According to Olin and Doris in their "Vicious Minds" paper (2013), it is "insofar as reliability is a condition on epistemic virtue" that we have reason to doubt that human organisms possess such virtue. The unreliability of agents follows from lack of 'stable' dispositions, due to our enormous vulnerability to irrelevant context effects. In order to formalize their challenge to VE they present the following dilemma involving a forced trade-off:

[V]irtue epistemologists encounter a dilemma: they can either formulate their theories to successfully accommodate the empirical challenge and lose the normative appeal derived from virtue ethics, or retain the normative appeal derived from virtue ethics and fall prey to the empirical challenge.<sup>57</sup>

We can call this the *trade-off problem*, because it is based on a thesis about a forced option between maintaining normative appeal and empirical adequacy. An embedded assumption I want to focus on is the *trade-off thesis*:

pressures on theory building in virtue epistemology are hydraulic: increase empirical adequacy at the expense of normative appeal, or increase normative appeal at the expense of empirical adequacy.<sup>58</sup>

Olin and Doris argue that virtue epistemology seems skepticism-inviting, but notice first that if the challenge to agent-reliability affects externalist and internalist epistemologies both, it may be that their own view is the one that invites skepticism. At least they say nothing to indicate what conditions *would* set a reasonable bar so that knowledge is not fleetingly rare. It looks like our modes of belief production on the vicious mind thesis are neither safe nor sensitive, so an analysis of knowledge involving tracking conditions or other sorts of counter-factuals may not pass muster either.<sup>59</sup> But more charitably we can take their

---

<sup>56</sup> Prinz, "The Normativity Challenge," 132.

<sup>57</sup> Lauren Olin and John Doris, "Vicious Minds," *Philosophical Studies* 168, 3 (2013): 665-692.

<sup>58</sup> Olin and Doris, "Vicious Minds," 32.

<sup>59</sup> It is hard to separate these issues of skepticism and anti-skepticism from concern with epistemic luck, which the authors do not mention; the distinction between Robust (RVE) an Anti-luck virtue epistemologies (ALVE) is quite as pertinent as the distinction between virtue reliabilists and responsibilists (and maybe more so, if the latter is largely a matter of emphasis and any VE must account for both knowledge-constitutive and auxiliary virtues). But luck isn't treated in Olin and Doris's paper. Since I am interested in virtue-theoretic responses to skepticism, I would argue for the advantages that allowing ALVE's independent anti-luck

concerns as focusing on credit theories, and more specifically on problems with an *aretaic* condition, including its best-known formulations in terms of a dispositional condition (Sosa) or an explanatory ‘because’ (Greco; Zagzebski).

Like Alfano, Olin and Doris do acknowledge my earlier comparison between the generality problem in epistemology and the situationist challenge to character-traits.<sup>60</sup> While they write that they “are sympathetic to the observation that the search for virtues and the search for reliable belief-forming process types may be partly coextensive,” they deny my argument that this attenuates the situationist challenge to the credit theory; they anyway find humans to be without many stable belief-forming process types. Since my claim of partial co-extension is granted but Olin and Doris do not comment on how I try to resolve this problem with a *Narrow-Broad Spectrum of Agency-Ascriptions* [Table 4], I will use this opportunity to provide a fuller development of how a properly modeled spectrum of agency-ascriptions assuages the concerns that Olin and Doris raise.

---

condition has over RVE attempts to have a single *aretaic* condition serve to deal with both the value problem *and* the epistemic luck problem. In addition to this key advantage I have elsewhere argue for, ALVE affords us a substantially smaller ‘empirical footprint’ than RVE. I cannot here explain my ‘smaller empirical footprint of anti-luck VE relative to Robust VE’ claim, but see Jesper Kallestrup and Duncan Pritchard, “The Power, and Limitations, of Virtue Epistemology,” in *Powers and Capacities in Philosophy: The New Aristotelianism*, eds. Ruth Groff and John Greco (London: Routledge, 2012), 248–269 and “Robust Virtue Epistemology and Epistemic Anti-Individualism,” *Pacific Philosophical Quarterly* 93, 1 (2012): 84–103 on advantages of ALVE (sometimes called ‘Weak VE’) in accounting for veritic luck cases especially in regard to what they term *epistemic dependence*, and the implausibility of any version of VE that cannot account for it. Pritchard also recognizes a spectrum of weak and strong epistemic achievements, where in many instances one can meet the rubric for cognitive achievements pretty easily, while in other instances the agent must overcome a significant obstacle to cognitive success, or manifest high levels of cognitive skill. See Guy Axtell, “Felix Culpa: Luck in Ethics and Epistemology,” *Metaphilosophy* 34, 3 (2003): 331–352; and “Two for the Show: Anti-Luck and Virtue Epistemologies in Consonance,” *Synthese* 158, 3 (2007): 363–383, for my first forays into a version of ALVE, which I ironically proposed in responding to Pritchard’s early Robust Anti-luck epistemology (RAL), the position he took in this first book, *Epistemic Luck* (Oxford: Oxford University Press, 2005).

<sup>60</sup> Olin and Doris, “Vicious Minds,” note 16.

**Table 4: The Narrow – Broad Spectrum of Agency-Ascriptions<sup>61</sup>**

<i>Narrowly-typed Abilities (NTA)</i>	<i>Broadly-typed Abilities (BTA)</i>
<ul style="list-style-type: none"> <li>- <i>Low-level virtues (faculty virtues).</i> Dispositions construed as genetically-endowed abilities or cognitive capacities.</li> <li>- Best suited to evaluating the etiology of a single belief or narrow range of beliefs;</li> <li>- Their ascription is often keyed to doxastic justification of particular beliefs actually held by an agent.</li> <li>- Their ascription answers to the <i>Generality Problem</i> by fixing the “narrowest, content-neutral process that is operative in belief production” for an <i>actually held</i> belief.<sup>62</sup></li> <li>- The value of low-level virtues is transmitted directly to their products and only indirectly to the agents who have them.</li> </ul>	<ul style="list-style-type: none"> <li>- <i>High-level virtues (reflective virtues).</i> Best suited to tell about the agent’s abilities and practices in a certain domain/area.</li> <li>- Best suited to holistic evaluation of agents, including the quality of their activities of inquiry.</li> <li>- Their ascription is not primarily to doxastic justification of particular beliefs. Their ascription is not to the ideology of ‘processes,’ but to the axiology of inquiry or ‘ideal-types.’</li> <li>- Their ascription is primarily occurrent, describing personal habits or counter-factual states of motivation an agent who performs a certain act-type in the given situation (and ethically or intellectually virtuous agent) <i>would</i> have.</li> <li>- The value of high-level virtues attaches directly to their possessor but only tenuously to their products.</li> </ul>

To develop the role of this *N-B Spectrum* in responding to the dilemma, I first want to distinguish between two senses of the trade-off thesis that drives Olin and Doris’ dilemma, replying to each sense in turn. The *formal sense* of the trade-off thesis indicates a *general* trade-off between pursuing normative assessment and pursuing descriptive adequacy. The *substantive sense* makes as I understand it a stronger and more directed claim: Proponents of an *aretaic* or person-level ability

<sup>61</sup> See also Lepock “Unifying the Intellectual Virtues,” from which my present chart and that of Axtell, “Character Trait Ascriptions,” draws many key points. Thanks to Christopher for discussion of his distinction and chart.

<sup>62</sup> Alvin I. Goldman, *Epistemology and Cognition* (Cambridge: Harvard University Press, 1988), 363.



condition face an especially difficult or dire version of this problem (relative to other (unspecified) potentially successful anti-skeptical epistemologies). They force this problem because whenever an *aretaic* condition is formulated so as to be empirically adequate (psychologically realizable) it will turn out to lack normative appeal, and whenever such a condition is formulated to have strong normative appeal, it will turn out to be empirically inadequate (psychologically unrealizable).

The *formal sense* of the trade-off thesis is a general warning about conflating explanatory and normative posits; if this is correct, it is largely a matter of logic. In response to this *formal sense* of the 'trade-off' problem, the *Narrow-Broad Spectrum* distinguishes different ways of ascribing epistemic agency to people. This *formal sense* makes no great worry for normative epistemologies because it is already acknowledged by the distinctions that the spectrum uses to avoid conflation that otherwise occur. Allowing that agency ascriptions serve a *variety of purposes*, as our spectrum does, means that what has the most empirical content *need not* have much normative appeal, and what has the most normative appeal *need not* have much empirical skin in the game.

The *formal sense* of the trade-off thesis locates it on the familiar grounds of the *normative-descriptive gap*. In so saying, however, familiar resources for responding also come to mind. The dilemma reflects the much-discussed and troublesome relationship between the normative appeal of unbounded rationality, with standards that ignore time, information, and computation of agents, and bounded rationality, where the scientific image, the real world, rushes back in. That we will be forced to judge humans as broadly irrational if we judge rationality by an 'Enlightenment picture' of unbounded agency when this is untrue of our actual way of being in the world, has been a common argument in the great rationality debate.<sup>63</sup> This is a rubbing point between N-theories and ecological rationality. But virtue epistemologies are not committed to any particular view on these broad issues. For them the *formal sense* of our dilemma just seems to be this well-known problem, re-directed towards conditions on knowledge rather than on rationality. There always is and needs to be a contrast of performance with competence, but standards of competence that arise entirely independently of psychology from ideals of pure logic or ideal agency, may lead us to doubt human competence universally.

Responding to the *substantive sense* of the 'trade-off' thesis is where our real work lies. Here there is more than a general warning, but a real risk of 'costs,'

---

<sup>63</sup> See David Matheson, "Bounded Rationality, Epistemic Externalism and the Enlightenment Picture of Cognitive Virtue," in *Contemporary Debates in Cognitive Science*, ed. Robert J. Stainton (Malden: Blackwell, 2006), 134-144.

in the present case to a credit theory of knowing. If we cannot say how and why *this theory* can avoid such conflation, or explain what legitimate epistemic concerns broadly-typed reflective virtues serve in an analysis of knowledge except to be causal – explanatory posits, then the theory looks to be in trouble. This requires a more detailed exposition of our *spectrum*. Responding to the substantive sense of the trade-off thesis requires making several further distinctions, including especially that between epistemological concerns with the *etiology* of belief (also called doxastic justification or *ex post* justification), on the one hand, and epistemological *axiology* together with norms of agential rationality/responsibility on the other.

At the Narrow end of the spectrum lie ascriptions of particular mental faculties and strategies, and at the Broad end lie more ‘occurrent’ ascriptions associated with agency under normal or ideal circumstances. It is only NT traits that directly address the generality problem, and hence doxastic justification, the core question of an account of knowing on all but purely internalist epistemologies.<sup>64</sup> This point basically follows from the normal reliabilist line on the generality problem, which is that the pertinent process-type to try to specify for doxastic justification is the “narrowest, content-neutral process that is operative in belief production.”<sup>65</sup> NT traits are best ascribed to a particular belief of an agent, to reconstruct its reliable etiology in a way that excludes all but modally remote error-possibilities. Agency ascriptions serving this etiological function, we can see, simply don’t require much normative ‘appeal,’ at least on the fairly minimal account of credit we discussed in connection with DPT in section 2.

On the other hand, the Broad end of the spectrum is directly concerned with praise; it is high praise, but does not in the same way entail epistemic credit, since there are plenty of ways to be personally justified (to be both synchronically rational as the evidentialist demands, and diachronically rational in our zetetic activities as the virtues would guide us to be), yet still come out with a false belief, or with a true belief not related in the *right way* to the agent’s ability, as in Gettier cases. Thus, if by attacking the ‘empirical adequacy’ of BT attributions Olin and Doris are assuming that such attributions always imply some robust disposition, this is simply mistaken. BT (broadly-typed) virtues are typically ascribed in assessment of an agent’s actions-at-inquiry – what I call *zetetic* activities – rather than in order to credit an agent for coming to have a *particular* a non-luckily true

---

<sup>64</sup> I do agree with some criticisms of a methodological individualist conception of knowledge, and I take VE to have strong overlaps with social epistemologists.

<sup>65</sup> See Goldman, *Epistemology and Cognition*, 363.

belief.<sup>66</sup> The ‘normative appeal’ of agency ascriptions and descriptions of the BT kind, I will need to argue, does not require them to have much empirical exposure. Since broad reflective virtues do not consist in a single neurological state of a person, that person’s behavior may be our best or only overt reason for ascribing such a dispositional property to them. The personal praise for broadly-typed virtues can be dispositional or occurrent. Occurrent attributions associate an agent with actions a virtuous or vicious agent *would* perform, or with motives they *would* have. We can call them occurrent because these kinds of counterfactuals demand only conformity with norms, and on the present view, are substantially weaker than causal-explanatory attributions.

Here we bump up against one of the major differences between ‘inquiry-focused’ VE and ‘classical’ responsibilism. I agree entirely with Olin and Doris’ point that ‘mixed’ accounts need to be careful what reliability and responsibility conditions are accepted and how they are framed. Inquiry-focused VE argues that intellectual virtues thought of as character-traits make us good at inquiry, but they contribute to a formal account of knowledge only in indirect ways. Hookway rightly notes that “Virtues regulate inquiries and deliberations, and only indirectly regulate beliefs.”<sup>67</sup> Since habits of mind acknowledged as intellectual virtues are rather abstract and complex traits, they accordingly have “a broad variety of possible manifestations in the intellectual activities.”<sup>68</sup> The motivations of reflectively virtuous inquirers normally do not regularly predict anything very specific by way of beliefs or activities.

It is clear that occurrent attributions needn’t be explanatory in the way that dispositional ones are, that the normativity of epistemic assessment and doxastic guidance differ substantially, and that the third-personal concerns with the etiology of belief (that the project of analysis requires) should be clearly distinguished from the first-personal concerns of the inquirer, which norm to what her epistemic community identifies and values as theoretical and/or personal virtues. What is less well-recognized is that virtue theories in ethics and epistemology can easily accommodate a narrow/broad, or causal-explanatory/normative spectrum of agency ascriptions, where posits at the

---

<sup>66</sup> Alfano, *Character as Moral Fiction*, 171, allows responsibilisms that “stand as a purely normative theory” without ‘explanatory-cum-evaluative’ ambitions. But this is in tension with claims he also makes that even Autonomous VE as Baehr defines it has conditions of realizability for virtues that it cannot meet. This may be a problem of semantics over what is designated by ‘character-traits.’

<sup>67</sup> Christopher Hookway “How to be a Virtue Epistemologist,” in *Intellectual Virtue*, eds. Michael DePaul and Linda Zagzebski (Oxford: Oxford University Press, 2003), 183-202, 197.

<sup>68</sup> Elke Brendel, “Function of Virtuous Dispositions,” 330.

different ends of the spectrum are recognized as serving substantially different albeit equally legitimate explanatory and evaluative functions.

Given this, it also follows that we should carefully distinguish not only dispositional from occurrent attribution, but overlapping with it, epistemic credit from praise. The Narrow end of our spectrum is concerned with epistemic credit, but not necessarily praise, because it may not be fitting to praise or blame someone for faculty virtues or 'automatic' responses, however successfully employed. Thus, if by 'normative appeal' Olin and Doris mean praise and censure, then its absence in NT ascriptions on the present account comes at no real 'expense.'

In most cases the concerns with doxastic responsibility and other aspects of BT virtues should be carried on apart from the project of analysis of knowledge. This is the general lesson of the failure of internalist evidentialism, and of intellectualist conceptions of mind, more generally.<sup>69</sup> Doxastic responsibility, like related concepts of rationality and personal justification, is taken as a generally necessary condition on knowledge only by more internalistically-oriented analyses. If one is going this route, the condition in question should be expressed negatively and occurrently, in the sense that meeting the condition merely means that the condition is met if we need *not* attribute to the agent motivations that an intellectually virtuous agent *would not* have, or actions or omissions that a virtuous agent *would not* perform or omit. Negative conditions on knowing (the agent's motivations were not vicious or her actions and omissions were not irresponsible, for instance) obviously are judged in the way occurrent ascriptions are, not by actual global-trait manifestation). And, of course, the more externalist versions of VE are still more cautious about a right motivational condition on knowledge. The focus for reliabilist VE is doxastic justification, which is what the generality problem relates to. For externalists the broad end of the spectrum just is not 'truth-linked' in the way that doxastic justification is.<sup>70</sup>

---

<sup>69</sup> See Axtell, "Recovering Responsibility."

<sup>70</sup> That responsibility conditions can easily be included in a formal analysis as *negative* rather than *positive* conditions is something overlooked when 'mixed' epistemology is treated as the epistemic situationist do, as always just adding another contentious generally necessary condition on knowing. Knowledge-attributions arguably range from animal ('brute'/externalist) to reflective (high-end/internalist) cognition. The beauty of negative conditions is that they allow sufficiency on knowledge in a plurality ways, and that they therefore fit with 'range concepts.' This is why situationism challenges Zagzebski's RVE more than ALVE. Since what classical responsibilists really care about is conditions on 'high-end' or specifically *reflective* knowledge, this is better put in a negative way of saying that the agent is 'not ill-motivated' or that her efforts at inquiry and evidence-gathering were 'not irresponsible,' etc., than by a general necessary condition on knowledge. My account thus accords with and adds support to

Some proponents of VE like Zagzebski who view the relationship between epistemology and ethics to be tighter than I do, may disagree. For those who try to reduce epistemic to ethical evaluation, or to treat personal responsibility as equivalent to doxastic justification, I reiterate the demand for a more general separation between the theory of knowledge and accounts of personal justification (synchronic and diachronic rationality).<sup>71</sup> We need to avoid the confusion of treating traits and states at the Broad end of the spectrum as doing the work of positive conditions on doxastic justification or knowledge, whose place is really with the generality problem and at the Narrow end. Critics of ‘classical responsibilism’ are probably right that trying to model epistemological normativity on ethical normativity (the neo-Aristotelian approach) contributes to these conflations.<sup>72</sup>

Our distinction between the formal and substantive senses in which the trade-off thesis can be construed has directly informed our response to Olin and Doris’ dilemma for VE. If theory construction and VE is ‘hydraulic,’ as they contend, then is it any *more* so than for other philosophical analyses of knowledge, and if so, why? Even granting that it is more so and there is a substantive worry about virtue epistemologies in particular, it surely is unfair not to allow virtue epistemologists the fluidity and flexibility of a spectrum of agency ascriptions in order to clarify the specific functions and levels of empirical exposure presupposed in different kinds of agency ascriptions. When we do, I

---

Kallestrup and Pritchard’s, which argues that ALVE fairs better on several dimensions, including empirical adequacy concerns, than does RVE.

<sup>71</sup> So consistent with my argument in “Recovering Responsibility,” that evidentialists like Feldman should accept the ‘separation,’ ‘divorce,’ or ‘two project’ proposal (that R. Foley and Anthony Booth persuasively argue for), I here place much the same demand upon *phronomic* virtue responsibilism. This result favors my more moderate stance of *zetetic* or inquiry-focused VE, which holds that reflective virtues are things that make us good at inquiry, and sometimes part of a rational reconstruction of belief acquisition or maintenance, but not for that reason a necessary condition on knowing.

<sup>72</sup> While I think avoiding these conflations is best addressed by inquiry-focused forms of VE, our account still agrees with Peter Samuelson and Ian Church when they write, “Presumably, any robust, full account of intellectual virtue will have to account for both cognitive faculty-virtues as well as character trait-virtues; whether one does this along agent-reliabilist lines or neo-Aristotelian lines could be, to some extent, a matter of emphasis. For a person can hold a belief more strongly (or weakly) than warranted due to biases inherent in one’s cognitive systems, or due to some lack of character, just as a person can exhibit virtuous knowing via the proper functioning of one’s cognitive system or through the exercise of a virtuous character.” (Peter Samuelson and Ian Church, “When Cognition Turns Vicious: Heuristics and Biases in Light of Virtue Epistemology,” *Philosophical Psychology* 28, 8 (2014): 1095-1113.)

believe that VE has advantages over other epistemologies in helping us address the *Normative-Descriptive Gap* problem, which is just what philosophers and psychologists working together should want.

A continuum or spectrum of agency-ascriptions helps to clarify what normative and explanatory concerns we have with different cases and different kinds of agency. But nothing we have said about agency ascriptions as they relate to knowledge possession implies the unreality of global and robust reflective virtues, or the idea that they are purely descriptive and never part of salient explanations for why we know. The reality of BT character – traits needn't be called into question by acknowledging the spectrum's range covering both dispositional and occurrent attributions. Being responsible is a great way to achieve epistemic goods like true belief, knowledge, and understanding. Indeed, without a lot of luck, this may be my only way, or at least my best shot at it. And the habits of responsible inquiry that I display today in coming to hold a true belief will, when an assessor retrospects on it tomorrow, be part of that belief's reconstructed etiology. But these contributions of responsibility to knowing still should not lead us to confuse etiology and axiology, or to think of personal justification (synchronic or diachronic) as guaranteeing these epistemic goods. Believing truly rather than falsely does not follow from my having unquestionable motives and giving my best cognitive effort. That is why Zagzebski's "because" condition is too strong, and why conforming to norms of motivation that a virtuous agent would have, or performing the inquiries a virtuous agent would perform, is generally sufficient to meet Broad-end norms associated with epistemic responsibility.

To conclude, these concerns with the formal and substantive senses of the 'trade-off' thesis are serious ones for virtue epistemologists, Olin and Doris are right to contend. But tying back to Section 2, DPT on my view also has the philosophical implication of supporting both the reality and value of global reflective virtues, moderating what can be claimed on empirical grounds about the modularity, or localness of character-traits, and the lack of robustness of traits of intellectual character. From a normative perspective, optimizing coordination of T1 and T2 within our natural limits is of crucial philosophical and pedagogical concern, especially since the *parallel* nature of T1 and T2 means they not only cooperate, but also both routinely operate at the same time and quite often compete in determining an agent's cognitive or ethical judgment.<sup>73</sup> These latter

---

<sup>73</sup> Compare Lisa Grover's argument that "We should accept the psychological reality of narrow, localized character traits, while retaining the thick evaluative discourse required by virtue ethics....Thick global concepts are necessary for a theory of localized character traits and

facts about how we process make the metacognitive prowess that comes with acquiring rational thinking dispositions *more* necessary and more efficacious than they appear to be on either situationist or automaticity ('System 1') theory. Habituating ourselves to rational thinking dispositions remains perhaps the most powerful tool within our adaptive toolbox.<sup>74</sup>

---

situation management to make sense. Without evaluative integration of different local traits under thick evaluative concepts we cannot identify which local traits to develop, and which situations to seek out, or avoid." (Lisa Grover, "The Evaluative Integration of Local Character Traits," *Journal of Value Inquiry* 46 (2012): 25-37, 36.

<sup>74</sup> Thanks to Mark Alfano, J. Adam Carter, Abrol Fairweather, Lauren Olin, John Doris, Anthony Booth, John Kihlstrom, Christopher Lepock, Christian Miller, and Holly Smith for comments and discussion. Lauren in particular provided a thoughtful and detailed set of comments on an early draft.





# A NEW RESPONSE TO THE NEW EVIL DEMON PROBLEM

Umut BAYSAN

**ABSTRACT:** The New Evil Demon Problem is meant to show that reliabilism about epistemic justification is incompatible with the intuitive idea that the external-world beliefs of a subject who is the victim of a Cartesian demon could be epistemically justified. Here, I present a new argument that such beliefs can be justified on reliabilism. Whereas others have argued for this conclusion by making some alterations in the formulation of reliabilism, I argue that, as far as the said problem is concerned, such alterations are redundant. No reliabilist should fear the demon.

**KEYWORDS:** dispositions, justification, The New Evil Demon Problem, reliabilism

The New Evil Demon Problem, presented by Cohen,<sup>1</sup> is meant to show that reliabilism of the sort that was defended by Goldman<sup>2</sup> is incompatible with the intuitive idea that the external-world beliefs of a subject who is the victim of a Cartesian demon *could* be epistemically justified. The original argument goes as follows:

- (1) If reliabilism is true, no external-world belief of a victim of an evil demon could be justified.
- (2) Some external-world beliefs of the victims of an evil demon could be justified.
- (3) Therefore, reliabilism is false.

One might think that there can't be much to add to the debate over the New Evil Demon Problem after more than thirty years of discussion. Nevertheless, there remains a *prima facie* plausible solution to this problem which hasn't been quite stated. In what follows, I shall present this solution.

As formulated as an objection to reliabilism, the argument that is sketched above takes reliabilism to be the view that a belief is justified if and only if it is formed as a result of a reliable belief-forming process. I shall call this version of reliabilism *crude reliabilism*. Just to restate, (3) holds that crude reliabilism is false.

---

<sup>1</sup> Stewart Cohen, "Justification and Truth," *Philosophical Studies* 46 (1984): 279-296.

<sup>2</sup> Alvin Goldman, "What Is Justified Belief?" in *Justification and Knowledge*, ed. G. Pappas (Ithaca: Cornell University Press, 1979), 1-23.

The reason for thinking that (1) is true is the following. Most (if not all) external-world beliefs of a victim of an evil demon are false because in demon worlds radical sceptical hypotheses are true: in a demon world, either there is no external world or the external world is radically different from the way it appears. Given the high frequency of false beliefs, the belief-forming processes of the habitants of demon worlds cannot be reliable; hence their beliefs cannot be epistemically justified. Or so the objector thinks.

The intuition that supports (2), which I shall call *the fairness intuition*, is that in some cases, victims of an evil demon might be doing all the right things in order to hold true external-world beliefs. When they believe that there are trees and cats in their surroundings, they do so because they undergo perceptual experiences that are subjectively indistinguishable from those experiences which are truly caused by real trees and cats. In fact, a victim of an evil demon could be an *epistemic counterpart* of you, and we might suppose that you have mostly epistemically justified beliefs. Here, I take an epistemic counterpart of S at  $t_1$  to be someone whose beliefs are, as far as their narrow contents are concerned, type-identical with the beliefs of S at  $t_1$ , and are furthermore held for the same subjectively accessible reasons as those of S at  $t_1$ . To illustrate: you believe at 11am today that it will rain; your reason for holding this belief is that you have a memory of the weather forecast reporting that it would rain. Your epistemic counterpart (as far as your temporal part at 11am today is concerned) believes that it will rain, and her reason for holding this belief is that she has a memory of the weather forecast reporting that it would rain. And it may be the case that whereas your belief is true, your epistemic counterpart's belief is false (or *vice versa*). The fairness intuition, I think correctly, suggests that if your beliefs are mostly justified, then your epistemic counterparts' beliefs should be mostly justified too. Assuming that you and your epistemic counterpart have the same reasons for holding same beliefs – you had the same perceptual experiences, have used the same inference rules, and so on – it is only *fair* to expect that your beliefs are justified if and only if your epistemic counterparts' beliefs are justified.

Let me just briefly state what I will *not* argue for. I will *not* argue that crude reliabilism can be weakened, or relativised, or made indexical, in order to accommodate the fairness intuition. Strategies along those lines have been endorsed by others,<sup>3</sup> and they do have their own virtues. But I believe that it is worth noting that such routes are *redundant*, at least as far as the New Evil Demon

---

<sup>3</sup> Alvin Goldman, *Epistemology of Cognition* (Cambridge: Harvard University Press, 1986) and Juan Comesaña, "The Diagonal and the Demon," *Philosophical Studies* 110 (2002): 249-266.

Problem is concerned. Crude reliabilism, without any further qualification, can accommodate the fairness intuition; we can formulate epistemic justification as the reliability of belief-forming processes, and still hold that our demonic epistemic counterparts' beliefs can be justified.

The key is to recognise that 'reliable' is a *dispositional* concept and, arguably, *reliability* is a dispositional property – insofar as it is a real property and there are dispositional properties. If one has problems with the idea of dispositional properties, most of what I will say can be understood in a non-dispositionalist framework. Take a true dispositional expression: "This vase is fragile." Why is this statement true? A full-blown realist about dispositional properties would say it is true *because* the vase that the "the vase" refers to is a bearer of the dispositional property of *being fragile*. Someone who is sceptical about dispositional properties, however, would say that the truth of this expression consists in the fact that the vase in question has some non-dispositional properties such that having those properties in the right circumstances makes it the case that the vase behaves in a fragile manner. The upshot is this: one needn't be a full-blown realist about dispositional properties in order to make sense of dispositional expressions.

Now consider *reliability* as a dispositional property. Take a supposedly true dispositional expression, such as "Lily is reliable." Whereas a realist about dispositionalist properties would say that this expression is true in virtue of the fact that Lily instantiates a dispositional property, namely *reliability*, an anti-realist about dispositional properties can still give a non-dispositional truthmaker about Lily for the said expression. I don't really want to be committed to any view about the reality or fundamentality of dispositional properties, but the points that I will make are easier to express with the resources of a dispositionalist view, so I will treat *reliability* as a dispositional property.

Many *sorts* of things can be reliable, and likewise, unreliable: people, machines, newspapers, weather, Wi-Fi signals, belief-forming processes, so on and so forth. When I say that Lily is reliable, arguably, I am not referring to the very same property of *reliability* that I refer to when I say that the Wi-Fi signal is reliable. A person's reliability consists in her disposition to tell the truth (or what she takes to be the truth) and keep her promises in the right circumstances. When the circumstances are not right, however, a reliable person might be forced to lie, or break a promise.

A belief-forming process's reliability consists in something quite different. Whereas the reliability of Lily is manifested in her telling the truth in the right circumstances, the reliability of a belief-forming process is manifested in the fact

that beliefs that are formed as a result of that process are mostly true, again, in the right circumstances. A reliable belief-forming process is disposed to produce true beliefs. That is, the manifestation of the dispositional property *reliability* attributed to a belief-forming process is the *truth* of the belief that is formed. Unreliable belief-forming processes, such as wishful thinking, aren't disposed to produce true beliefs. Occasionally, the beliefs that are formed as a result of wishful thinking may turn out to be true. But this is not different from the occasional breaking of non-fragile vases. Such occasional breakings don't have to be miraculous. Vehicles like the Popemobile and the Batmobile have windows made of non-fragile glass, yet presumably they couldn't stay intact after an atomic bomb explosion. Our standards for non-fragility are not so high that only absolutely unbreakable things can be deemed non-fragile.

Although the reliability of a person and the reliability of a belief-forming process might be different properties, the rules of the application of the predicate 'is reliable' to people and to belief-forming processes are similar in an interesting way. The similarity lies in the fact that *one can be a bearer of a dispositional property without ever manifesting the disposition in question*. Strictly speaking, it is possible for a reliable person to lie *at all times*. Admittedly, this sounds very odd; nevertheless it is true. It belongs to the concept of 'disposition' that dispositions needn't be manifested in order to be instantiated. There are fragile vases which are never broken, simply because they have never been struck. So, the following is a perfectly possible state of affairs:

- (i) *a* is fragile; *a* is not struck; *a* doesn't break.

But more strangely, there *could be* fragile vases that are never broken, despite being struck and dropped multiple times. Think of the case of the sorcerer who is the guardian of a fragile vase.<sup>4</sup> Every time the vase is struck, the sorcerer casts a spell on it so that it resists the strike. The vase in question still counts as fragile; if the sorcerer weren't guarding it, it would have manifested its fragility. (Note that this is true non-vacuously: the sorcerer is only contingently protecting the vase.) So, the following is a perfectly possible state of affairs too:

- (ii) *a* is fragile; *a* is struck many times; *a* doesn't break.

Moving on from fragile vases to reliable people: consider the case of Lily. Lily is disposed to tell the truth, but for some reason, at every single attempt, she fails to do so. Maybe her actions are being manipulated by the Purple Man, who is a master of mind-control. Lily wants to tell the truth; she genuinely intends to do

---

<sup>4</sup> David Lewis, "Finkish Dispositions," *Philosophical Quarterly* 47 (1997): 143-158.

so, but every time she speaks, she lies. She, I stipulate, is still reliable, but she is not manifesting her reliability, because she is being controlled by the Purple Man. If the Purple Man weren't manipulating her actions, Lily would have told the truth. (Again, this is true non-vacuously: the Purple Man is only contingently manipulating Lily's decisions.) So, the following is a possible state of affairs:

- (iii) *a* is a reliable person; *a* is asked if *P* is true; *a* knows that *P* is true; *a* says that *P* is false; this happens systematically.

I hope I have convinced you that (ii) and (iii) are possible states of affairs. If you still have doubts, remember that dispositions require right circumstances in order to be manifested in the right way. By introducing sorcerers and mind-controlling supervillains, we are departing from right circumstances.

Now, beliefs. A belief forming-process may be disposed to produce true beliefs, but for whatever reason, at every attempt, it may fail to do so. As I hope is clear from the discussion so far, all we need to do is depart from the right circumstances. In a demon world, what is happening is exactly this. The deeds of the evil demon change the circumstances so the belief-forming processes, however reliable they are, are not manifesting their reliabilities. So, the following is also a perfectly possible state of affairs:

- (iv) *a* is a reliable belief-forming process; *a* is exercised; *a* doesn't produce true beliefs; this happens systematically.

Now if (iv) is really a possible states of affairs, premise (1) of the argument above is false: one can be a crude reliabilist about epistemic justification and still hold that external-world beliefs in a demon world can be epistemically justified. If all this is right, then it appears that crude reliabilism doesn't have to be weakened or relativised in order to accommodate the fairness intuition. What needs to be done is to recognise that *reliability* is a dispositional property and remember that dispositions can be held without ever being manifested.

If I am right, crude reliabilism, a version of reliabilism which has been abandoned partly due to worries about the New Evil Demon Problem, actually has the resources to deal with this problem. I showed a hitherto unexplored and *prima facie* plausible logical space where both crude reliabilism and the fairness intuition are true.<sup>5</sup>

---

<sup>5</sup> **Acknowledgments.** Many thanks to Robert Cowan and Martin Smith for discussion and comments on a previous version of the paper. This publication was made possible through the support of a grant from the John Templeton Foundation. The opinions expressed in this publication are those of the author and do not necessarily reflect the views of the John Templeton Foundation.



# NON-PICKWICKIAN BELIEF AND ‘THE GETTIER PROBLEM’

John BIRO

ABSTRACT: That in Gettier's alleged counterexamples to the traditional analysis of knowledge as justified true belief the belief condition is satisfied has rarely been questioned. Yet there is reason to doubt that a rational person would come to believe what Gettier's protagonists are said to believe in the way they are said to have come to believe it. If they would not, the examples are not counter-examples to the traditional analysis. I go on to discuss a number of examples inspired by Gettier's and argue that they, too, fail to be counter-examples either for reasons similar to those I have urged or because it is not clear that their subject does not know.

KEYWORDS: Edmund Gettier, knowledge, belief

I. Few things are more widely agreed upon by philosophers today than that the traditional analysis of knowledge as justified true belief was dealt a damaging, perhaps fatal, blow by Edmund Gettier in his famous paper published more than fifty years ago.<sup>1</sup> Since that time, most discussion of the topic has centered on how to repair the analysis either by beefing up the justification condition or by adding a fourth one. Few have questioned Gettier's claim that in his alleged counter-examples to the analysis all three conditions claimed to be necessary are satisfied. That the truth condition is, cannot be doubted: that is a matter of stipulation. There has been some debate, though not much, and usually only with respect to the first example, whether Gettier's protagonist really satisfies the justification

---

<sup>1</sup> Edmund Gettier, “Is Justified True Belief Knowledge?” *Analysis* 23 (1963): 121-123. Examples of the many confident pronouncements to this effect: “Gettier's counter-examples leave the justified-true-belief theory stone dead.” (David Owens, *Reason without Freedom: The Problem of Epistemic Normativity* (London: Routledge, 2000), 41); “Gettier described two cases that decisively refute the analyses of knowledge as justified true belief.” (Mathias Steup, “The Analysis of Knowledge,” in *The Stanford Encyclopedia of Philosophy* (Summer 2012), ed. Edward N. Zalta, URL=<http://plato.stanford.edu/archives/sum2012/entries/knowledge-analysis>). As Williamson, who regards Gettier's examples as paradigms of thought experiment, remarks, “...his refutation of the justified true belief analysis was accepted almost overnight by the community of analytic epistemologists.” (Timothy Williamson, *The Philosophy of Philosophy* (Malden: Blackwell, 2007), 180) I shall consider some of Williamson's own Gettier-style cases below.

condition.<sup>2</sup> Even fewer have questioned – and, again, only in connection with the first example – whether the belief condition is satisfied.<sup>3</sup> Indeed, even the possibility of defending the traditional analysis in this way has been denied. According to Floridi, one can try to defend it *only*

by strengthening/modifying the only flexible feature of the account, namely the justification condition; or by adding at least one more condition that would prevent the Gettier-ization of the required justified true beliefs or, alternatively, allow their de-Gettierization; or by combining (a) and (b). *No other general strategies are available* (my emphasis).<sup>4</sup>

Here I aim to show that in neither of Gettier's cases are both the belief condition and the justification condition satisfied and thus that the cases do not constitute counter-examples to the traditional analysis. I shall also discuss a number of examples modelled on those of Gettier's to show that they fail for similar reasons.

I take it that the following characterization of a standard 'Gettier' case would be widely accepted: S believes that *p*, S is justified in believing that *p*, *p* is true, but S's justification for believing that *p* is rooted not in the fact that makes *p* true but in some false proposition S is justified in believing from which *p* follows. It is then claimed that it is intuitively clear that S does not know *p*.<sup>5</sup> While Gettier's claim that in the cases he describes the belief condition is satisfied has gone virtually unquestioned, I think there is reason to question it. I aim to show

---

<sup>2</sup> Christopher New, "Some Implications of 'Someone,'" *Analysis* 26, 2 (1965): 62-64 and "Someone' Renewed," *Analysis* 28, 3 (1968): 109-112; Charles Pailthorp, "Knowledge as Justified, True Belief," *The Review of Metaphysics* 23, 1 (1969): 25-47; Irving Thalberg, "In Defense of Justified True Belief," *The Journal of Philosophy* 66 (1969): 794-803.

<sup>3</sup> Christoph Schmidt-Petri, "Is Gettier's First Example Flawed?" in *Proceedings of the 26th International Wittgenstein Symposium. Austrian Ludwig Wittgenstein Society* (Kirchberg am Wechsel, 2003), 317-319; Benoit Gaultier, "An Argument Against the Possibility of Gettiered Beliefs," *Logos & Episteme* V, 3 (1914): 265-272.

<sup>4</sup> Luciano Floridi, "On the Logical Unsolvability of the Gettier Problem," *Synthese* 142 (2004): 62

<sup>5</sup> In what I am calling a *standard* case, the (supposed) belief that turns out to be fortuitously true is a false but justified belief. Other examples that supposedly show the inadequacy of the traditional analysis, such as Lehrer's Grabit, Harman's assassination and Goldman's fake-barn cases, are often lumped together with Gettier's, even though they do not exhibit the standard pattern and – perhaps for that reason – are more controversial, with intuitions dividing on whether their subject knows or not. (Keith Lehrer and Thomas Paxson, Jr., "Knowledge: Undefeated Justified True Belief," *Journal of Philosophy* 66, 8 (1969): 225-237; Gilbert Harman, *Thought* (Princeton: Princeton University Press, 1973); Alvin Goldman, "Discrimination and Perceptual Knowledge," *Journal of Philosophy* 73, 20 (1976): 771-791. Goldman tells us that the example comes from Carl Ginet.) I shall say something about such cases later in the paper.



that in neither of the cases Gettier describes does Smith believe the proposition Gettier claims he does. If I am right, the cases Gettier describes make no trouble for the traditional analysis.<sup>6,7</sup>

In the first challenge I know of to the claim that in the first case Smith does believe the proposition that turns out to be true, Schmidt-Petri argues that the definite description ('the man who will get the job') in the sentence that is supposed to express the proposition for which Smith has good evidence ("Jones is the man who will get the job, and Jones has ten coins in his pocket" – Gettier's (d)) is used referentially, whereas the same definite description in the sentence expressing the proposition Smith supposedly infers from it ("The man who will get the job has ten coins in his pocket" – Gettier's (e)) it is clearly used attributively. But surely, should Smith utter (e), he would be expressing a belief about Jones and Jones only.<sup>8</sup> It is, in general, not difficult to recognize which use is in play. When

---

<sup>6</sup> The claim that the belief condition is not satisfied must not, of course, be confused with the claim that it is not a necessary condition, as Radford suggested. (Colin Radford, "Knowledge by Examples," *Analysis* 27, 1 (1966): 1-11.) I agree that it is not, though not for the same reasons. But whether we are right in this has no bearing on whether in Gettier's examples the condition is satisfied. If it is not, the examples are not counter-examples to the sufficiency of the traditional analysis. Reasons for thinking that believing is not a necessary condition of knowing, either, may be found in J.L. Austin, "Other Minds," *Proceedings of the Aristotelian Society*, Supplementary Volume 20 (1946):171 and Zeno Vendler, *Res Cogitans* (Ithaca, NY: Cornell University Press, 1972), Ch.5; reasons for thinking that being justified is not a necessary condition of knowing, in Crispin Sartwell, "Knowledge is Merely True Belief," *American Philosophical Quarterly* 28 (1991): 157-65).

<sup>7</sup> This is, of course, not enough to show that there are no genuine Gettier-style counterexamples to the traditional analysis. Stopped-clock and fake-barns cases may be thought plausible candidates. The difference is that in these the belief condition is – arguably – satisfied, whereas, as I shall argue, there is reason to think that in the former it is not. Some think that in one or the other or both of these cases the subject does have knowledge. (Stephen Hetherington, "Knowing Failably," *Journal of Philosophy* 96 (1999), 565-587; Igor Douven, "A Contextualist Solution to the Gettier Problem," *Grazer Philosophische Studien* 69, 1 (2005): 207-228; William G. Lycan, "On the Gettier Problem Problem," in *Epistemology Futures*, ed. Stephen Hetherington (Oxford: Oxford University Press, 2006), 148-168). Heathcote thinks this with respect to the fake-barn case but views the stopped-clock case as a "classical Gettier counter-example." (Adrian Heathcote, "Gettier and the Stopped Clock," *Analysis* 72, 2 (2012): 309). I can be argued, though, that in that case the belief and the justification conditions fail to be jointly satisfied. (John Biro, "Showing the Time," *Analysis* 73, 1 (2013): 57-62.) I discuss these cases further in section IV.

<sup>8</sup> More recently, Gaultier has offered a different reason for denying that Smith believes the second proposition (Gaultier, "An Argument"). Because while Gaultier confines his discussion to Gettier's first case, his argument is general enough to be easily extended to the second, I reserve discussion of it until after I have considered the latter.

Dr. Johnson says “The man who is tired of London is tired of life” (as he is sometimes reported, inaccurately, to have done), we do not take him to have learned that a particular man was tired of London and also tired of life. If we did, the utterance would obviously lose the interest it now has for us. What Johnson actually said was “When a man is tired of London, he is tired of life,” which is clearly not about any particular man. (Johnson – of all people – would not have said what he is reputed to have said, knowing that it could be interpreted to mean what he did not.) He could have said “He who is tired of London is tired of life,” much as we say “He who hesitates is lost.” But we cannot substitute ‘he who’ for ‘the man who’ in “I am the man who broke the bank at Monte Carlo.” Yet this is, in effect, what Gettier is asking us to do in suggesting that we can infer the belief that is made true by Smith’s having ten coins in his pocket from our false belief that Jones will get the job and has ten coins in his pocket.

Another test of whether a definite description is being used referentially or not is adding ‘namely’ followed by a name to see if that yields something the speaker or thinker can be plausibly supposed to believe.<sup>9</sup> Adding ‘namely, Jones’ in (e) does. But if that is what Smith believes, *his* getting the job does not make his

---

<sup>9</sup> The ‘namely’ test was suggested by New (“Some Implications”) in connection with Lehrer’s Nogot/Havit example, as a way of distinguishing between the valid inference from “A is F, therefore someone, namely, A is F” and the invalid one “A is F, therefore, someone who may not be A is F.” (Keith Lehrer, “The Gettier Problem and the Analysis of Knowledge,” in *Justification and Knowledge*, ed. George S. Pappas (Dordrecht: Kluwer/Reidel, 1979), 65–78.) New argues that the sentence “Someone is F” is ambiguous in that it can be used to “express either a statement to the effect that an *identifiable* someone is F or “a statement to the effect that an *unidentifiable* someone is F. This is clearly close to – and anticipates – Donnellan’s famous distinction (with the difference that on the latter the difference is not one of meaning but of use). (Keith Donnellan, “Reference and Definite Descriptions,” *The Philosophical Review* 75 (1966): 281–304). However, New’s interest is, as usual, in how overlooking the difference affects the claim that the justification condition is satisfied. So is Pailthorp’s and Thalberg’s in making similar complaints. Bernecker also thinks that the first example turns on ambiguity: “In Gettier’s first example Smith is said to have a true and justified belief to the effect that the man who will get the job has ten coins in his pocket. The proposition that Smith believes – the man who will get the job has ten coins in his pocket – is ambiguous. Smith takes the definite description to refer to Jones but it in fact picks out Smith. If the definite description refers to Jones, Smith’s belief turns out to be justified but false. If the definite description refers to Smith, the belief is true but unjustified. The example therefore fails to show that justified true belief is insufficient for knowledge.” (Sven Bernecker, “Keeping Track of the Gettier Problem,” *Pacific Philosophical Quarterly* 92, 2 (2011): 127–52). Note, again, the focus on justification and the assumption that the belief condition is satisfied. However, Bernecker and I (and Schmidt-Petri) do agree that there is no proposition such that all three conditions are satisfied with respect to it.

belief true. Adding 'namely, Smith' also yields a belief, but it is not one he can be plausibly supposed to have. The only remaining reading of the definite description in (e) is as attributive. If so, this blocks the supposed inference from (d) to (e). Not only would such an inference fail to transmit justification, but Smith, being rational (as we must obviously assume him to be) would not in fact make it. Being rational, he must believe that even though he has good evidence that Jones will get the job, that evidence does not warrant believing that whoever gets the job has ten coins in his pocket. But if Smith does not believe (e) understood attributively, the example shows nothing about the adequacy of the traditional analysis.

However, this objection to Gettier's first example focuses on what is in fact an accidental feature of it. Other so-called Gettier examples, including his own second, do not involve an ambiguity of the sort that infects the first example. And it is easy to amend even that one by re-phrasing it without the ambiguous definite description. "Someone will get the job and that person has ten coins in his pocket" follows from "Jones will get the job and he has ten coins in his pocket," and if Smith recognizes this, he will, it may be thought, surely infer the former from the latter. (But see below.) His justification for believing the second proposition then carries over to the first, and if that happens to be true in the way Gettier asks us to imagine, he has a justified true belief but, it seems, no knowledge.

Another version of the attempt to undermine the first example by appealing to an ambiguity is to invoke a distinction between what Heathcote calls the speaker meaning interpretation and the objective referent interpretations, respectively. He argues that "...when we disambiguate (e), we either get a justified false belief or we get a unjustified true belief – but in neither case do we get a justified true belief."<sup>10</sup> But this objection, too, appears to be blunted by eschewing the definite description.

---

<sup>10</sup> Adrian Heathcote, "Truthmaking and the Gettier Problem," in *Aspects of Knowing: Epistemological Essays*, ed. Stephen Hetherington (Oxford: Elsevier, 2006), 153. Heathcote concedes that Gettier's second example is immune to this objection. (See footnote 22 below.) A variant of the same strategy is deployed by Mizrahi, who, appealing to Kripke's notion of ambiguous designators argues that Gettier cases – Gettier's own and the other discussed here – are misleading in that, contrary to the usual understanding, they reveal a semantic mistake, rather than an epistemic failure. (Moti Mizrahi, "Why Gettier Cases Are Misleading," *Logos & Episteme* VII, 1 (2016): 31–44). What is not clear in Mizrahi's discussion is why he thinks that the semantic mistake does not carry in its train the epistemic shortcoming alleged by those who take Gettier cases to be counterexamples to the traditional analysis. The account offered here can give a natural explanation of this: if in saying or thinking that something has a certain property one has in mind something other than the thing of that actually has that property, one does not believe that the latter has the property. If so, the belief condition of the traditional analysis is not satisfied with respect to the proposition that turns out to be true. This account

In any case, neither Gettier's second example nor the many other Gettier-style examples inspired by him involve such ambiguities. Take Chisholm's well-known one, in which one mistakes a bush for a sheep, but one's belief that there is a sheep in the field still turns out to be true, since there is an unseen sheep behind a tree.<sup>11</sup> "There is a sheep in the field" contains no definite description. The same goes for Lehrer's Nogot/Havit case and for Turri's more recent Lamborghini case (both to be discussed below).<sup>12</sup>

There is, though, another way to object: a way that brings out that the right question to ask is not whether Smith would be justified in believing (e) but whether he would believe it in the first place. Grant that "X is F" entails that "Something is F" and that I, who believe that X is F, know this. Is this enough to lead me to form the belief that something is F, a belief that would be made true by Y's being F?

Another way to bring out the point may be to ask what proposition is expressed by the sentence "The man who will get the job has ten coins in his pocket." It is supposed to be one made true by Smith's getting the job and having ten coins in his pocket. If so, it must be the proposition – (e) – that whoever gets the job has ten coins in his pocket. That proposition is one about a relation between two properties. By contrast, (d) says only that they are both instantiated in one, possibly idiosyncratic, case.<sup>13</sup> What reason is there to think that the possibly accidental co-instantiation of two properties reveals some interesting relation between them?

---

also has the virtue of not requiring reliance on controversial views about meaning and reference, such as Grice's and Kripke's.

<sup>11</sup> Roderick Chisholm, *The Theory of Knowledge* (Englewood Cliffs: Prentice-Hall, 1977), 23

<sup>12</sup> Lehrer, "The Gettier Problem," 23; John Turri, "In Gettier's Wake," In *Epistemology: The Key Thinkers*, ed. Stephen Hetherington (London: Continuum, 2012), 214-229.

<sup>13</sup> It is, no doubt, possible for someone to infer (e) from (d), thinking that the latter follows from the former. But if (e) is taken to express a belief that is supposed to be made true by Smith's getting the job, that belief cannot be justified on the basis of (d). Someone who recognized this would not make the inference and would therefore fail the belief condition. Someone who failed to do so and made the inference would fail the justification condition. Either way, one of the conditions deemed necessary by the traditional analysis would be left unsatisfied. Recall that in Donnellan's example about the murderer of Smith, in the scenario illustrating the attributive use it is the nature of the crime that grounds the belief, whereas in the one illustrating the referential one it is something entirely independent of it. Nothing in the second suggests a relation between the relevant properties. (Suppose the crime not particularly grizzly, but the man in the dock behaving as Donnellan describes, and suppose him innocent and the real culprit insane.)

II. Of course, even if Gettier's first example fails to be a genuine Gettier case and thus a counter-example to the traditional analysis, his second may succeed.<sup>14</sup> Here we are asked to grant that S believes that either Jones owns a Ford or Brown is in Barcelona. (Gettier's (h)). He is said by Gettier to have 'constructed' (h), along with (g) that either Jones owns a Ford or Brown is in Boston and (i) that either Jones owns a Ford or Brown is in Brest-Litovsk. According to Gettier, Smith believes all these disjunctions, since he believes the first disjunct. No-one, to my knowledge, has suggested that we should not grant this. But I think there is reason not to. The assumption on which the claim that Smith believes these disjunctions rests is that what we may call belief tables mirror truth tables. This, I suggest, is an assumption we should not grant.

Is it really the case that anyone could believe (h) in a non-pickwickian sense? I take it that believing something in a non-pickwickian sense means being prepared to assert it (seriously, not pretending to, as one may in logic class). Being prepared to assert a disjunction implies that one is prepared to assert *either* one of the disjuncts on finding the other to be false. This may seem a surprising and unreasonable requirement. After all, is not the truth of one of the disjuncts sufficient for the truth of the disjunction? Smith is prepared to assert one of the disjuncts; why would his not being prepared to assert the other disjunct be a barrier to being prepared to assert the disjunction? The reason is not unlike the one we saw blocking a rational inference from (d) to (e) in Gettier's first example. Unless I *know* one of the disjuncts, I cannot, if I am rational, rule out the possibility that the one I believe (and think myself, perhaps rightly, justified in believing) is false. Believing that this is possible, I will be prepared to assert the disjunction only if I am prepared, on learning that one of the disjuncts, namely, the one I believed to be true, is false, to assert the other. But this is not so with Gettier's 'constructed' disjunctions.

The suggestion is that being prepared to assert seriously that  $p \vee q$  requires not only being prepared to assert either that  $p$  or that  $q$  but also being prepared to assert that  $(\neg p \rightarrow q) \wedge (\neg q \rightarrow p)$ . This is not quite to require that the disjunction be understood as exclusive, since the requirement is compatible with being also prepared to assert that  $p \wedge q$ . (One sometimes says: " $p \wedge q$ , but at least  $p \vee q$ .")

---

<sup>14</sup> There is considerable variation in what is counted as a Gettier case. In characterizing 'the Gettier problem,' Steup does not even discuss the first example. Schmidt-Petri, by contrast, claims that it is 'the Gettier example.' Pritchard describes the stopped-clock example as 'the paradigm Gettier-style counter-example' (Duncan Pritchard, *Epistemic Luck* (Oxford: Oxford University Press, 2014), 156) and uses it as his stalking horse, without even mentioning Gettier's own cases. Heathcote thinks that it, but not the fake-barn case, is 'in the classic Gettier mold.' (Heathcote, "Gettier and the Stopped Clock," 309)

What a serious assertion of " $p \vee q$ " is not compatible with is being prepared to assert *only* that  $p \wedge p \vee q$ , without being prepared to assert that if  $\neg p$ , then  $q$ . But, surely, that is Smith's situation: he is not prepared to assert for any of the disjunctions he is said to have constructed that if the first disjunct is false, the second is true. Were he prepared to assert this with respect to one of them, he would have to be prepared to do so with respect to all three (not to mention the indefinitely large number of others he could construct in the same way). Clearly, someone rational would not be prepared to do this.

With respect to belief, unlike with truth and even, perhaps, assertion, ordinary 'or' *is* to be interpreted as always standing for exclusive disjunction. While  $p \vee q$  is true even if both  $p$  and  $q$  are true (difficult enough to make a non-philosopher see at first), to say that someone believes that  $p$  or  $q$  (in a non-pickwickian sense), is not only not to say that he believes that  $p$  and also believes that  $q$ ; it is at least to imply that he believes only one of them. The correct way to report that someone believes both is to say that he believes that  $p$  and  $q$ . There is no analogue in ordinary belief talk of the inclusive 'v' of the truth table. Believing that  $p$  or  $q$  amounts to believing that one or the other is true but being undecided as to which that is. That is obviously different both from believing that  $p$ , as well as from believing that  $p$  and  $q$ .<sup>15</sup>

Even if this were not so generally, there is another reason why at least in Gettier's second example 'or' must be interpreted as exclusive. Presumably Smith would sign on to the conjunction of the three propositions he is said to have constructed only if he so took it, knowing that they could not all be true by way of both their respective disjuncts' being true. So, we need to take him as believing that either Jones owns a Ford or *else* he is in Barcelona. But then we must do the same with his other 'constructs' and take him to believe that either Jones owns a Ford or *else* he is in Boston, etc. Doing so would require that on learning that Jones does not own a Ford he believe – be prepared to assert – the conjunction of the second disjuncts of the constructed propositions.

Thus for Smith to believe (h), (g) and (i), he would have to believe that if Jones does not own a Ford, Brown is in Barcelona *and* if Jones does not own a Ford Brown is in Boston *and* if Jones does not own a Ford Brown is in Brest-Litovsk. He obviously does not believe all this. But if he is to satisfy the conditions of the traditional analysis with respect to (h), we must think of him as believing it. We

---

<sup>15</sup> It may be asked, can one not believe that  $p$  and possibly also  $q$ ? Yes, but so understood, all the 'constructed' propositions (and many more) are true, and Brown's actual whereabouts are irrelevant to their truth. In any case, the propositions Jones 'constructs' and supposedly believes are categorical, not modal.

must thus think that there is something about (h) to make him believe *it* and be justified in believing it, rather than any of the others he has constructed (or could construct). What could that be?

Suppose, further, that believing that Jones owns a Ford, Smith constructs (j) "Jones owns a Ford or the moon is made of green cheese" or (k) "Jones owns a Ford or 4 is prime." Should we think of him as seriously believing these? Even if, as we are supposing, he has good evidence for the first disjunct, he presumably realizes that it could be false. And he knows that  $p \vee q$  entails  $\neg p \rightarrow q$ . Can we think of him as seriously believing that if the first disjunct of (j) or (k) is false, their respective second disjuncts are true? (j) is unlikely to be made true by its second disjunct and (k) cannot be. Thus neither disjunction is something of which it can be said that Smith seriously believes it, hence neither satisfies the traditional analysis. It is no different with (h), which is 'constructed' in the same way.

In fact, Smith could 'construct' by addition any disjunctive proposition whatever and any of these could be made true by the truth of the added disjunct. Smith knows this. What is the difference between his believing the three he is said to believe and his not believing all the others? Just that those three happened to occur to him? Or should we say that he believes all of them? Neither choice strikes me as attractive.

III. Showing that Gettier's own examples are not genuine 'Gettier' cases and thus do not refute the traditional analysis is, of course, not enough to show that none of the many similar examples in the literature succeed where his fail. Obviously, I cannot review them all here. But a look at a few well-known ones may help bring out a general reason for thinking that they are all likely to be flawed in the same way. In Lehrer's Nogot/Havit case, the inference is by way of existential generalization. On the face of it, such an inference is unexceptionable. If Nogot, who is in the room, owns a Ford, it does follow that someone in the room owns a Ford. So, why balk at saying that someone who believes the first and infers the second from it believes the latter? Indeed, Lehrer himself does not, seeing the problem as one having to do with justification. He proposes a fourth condition, failure to satisfy which would "...have the effect of blocking... the transmission of justification" and thus underwriting the intuitively correct verdict that one would not know that someone in the room owned a Ford.<sup>16</sup> Thus, again, it is the

---

<sup>16</sup> Lehrer, "The Gettier Problem," 25. The details of how the fourth condition is supposed to do this are not important for present purposes. (As mentioned earlier, sometimes adding a fourth

justification condition that is seen as not satisfied, with the satisfaction of the belief condition assumed.

However, adding a fourth condition (or strengthening the justification condition) is conceding that, unimproved, the traditional analysis *is* inadequate. We are not forced into doing this if it can be shown that in this case and in similar ones the belief condition is not satisfied. And we can do this if we look more closely at whether the inference as described yields a seriously held, non-pickwickian, belief.

Here the inference has the form  $Fa \wedge Ga$ , therefore  $Ex (Fx \wedge Gx)$ , which is obviously valid, unlike the one from (d) to (e) in Gettier's first case. Yet even here it does not follow from one's believing that Nogot owns a Ford that one believes that someone or other in the room (that is, if not Nogot, then someone else) owns a Ford, which is what one would have to believe, if one were to believe something that Havit's owning a Ford could make it true. For one to have such a belief, one would have to believe that if Nogot does not own a Ford, someone else in the room does. In Lehrer's story, one has no reason to believe that. Thus, again, either one does not make the inference or, if one does, one is not justified in doing so.

As before, the fact that the truth of the premise entails the truth of the conclusion should not be taken to mean that believing the premise (however strong the evidence for it) and recognizing that it entails the conclusion suffices for someone rational to believe the latter (in a non-pickwickian sense). Believing it in that sense requires, as we saw in connection with Gettier's second case, that one be prepared to assert that if, despite one's evidence, Havit does not own a Ford, someone else in the room does. This is something our reasoner is presumably not prepared to assert, having no evidence that suggests it. Hence we have no reason to say that in this case the belief condition is satisfied and thus no reason to think that we have a genuine Gettier case.

In Chisholm's example, the presence of an unseen sheep does not make true my mistaken belief that what I am looking at is a sheep.<sup>17</sup> It makes the sentence

---

condition and beefing up the justification condition come to the same thing. This is the case here.)

<sup>17</sup> The example fails the analogue of the 'namely' test. City Slicker: "There is a bull in the field!" Farmer: "Where?" "There, by the creek!" "Nah, that's a cow." An unseen bull in the field would not make what City Slicker believes true, even though the proposition it would make true follows from the one he does believe. (Contrast: Farmer: "There is a bull in the field." City Slicker: "Where?" Farmer: "I don't know, but those cows are sure acting nervous.") Imagine that we are on safari, hoping to catch a glimpse of the rare and elusive grumpus. I whisper, excitedly, "There is one!" You, my guide, deflate me by saying, "No, that just looks like one, it is a common pumpus. But I know there is a grumpus somewhere in this area – I have seen its tracks." I am



"There is a sheep in the field" true, but, as is well known, belief is finer-grained than sentence meaning. (Consider, most obviously, sentences containing indexical terms.)

Williamson claims that the following example exhibits the same pattern:

A clever bookseller fakes evidence which appears to show conclusively that a particular book once belonged to Virginia Wolf; convinced, Orlando pays a considerable sum for the book. He has a justified false belief that this book of his once belonged to Virginia Wolf. On that basis alone, he forms the existential belief that he owns a book which once belonged to Virginia Wolf. The latter belief is in fact true, because another of his books in fact once belonged to her, although he does not associate that one with her in any way. Thus Orlando has a justified true belief that he owns a book that once belonged to Virginia Wolf, but he does not know that he owns a book which once belonged to Virginia Wolf.<sup>18</sup>

Williamson is right that the example follows a familiar pattern. It should not come as a surprise that it fails for a by now familiar reason: it does not pass the 'namely' test, as is shown by the fact that the object of Orlando's justified false belief is described by Williamson himself as '*this* book of his' (my italics), namely, the ringer. That belief is not made true by his unwitting ownership of another book that did belong to Virginia Wolf.

Williamson also offers what he calls a real-life Gettier example. Here he describes himself as apologizing to an unsuspecting audience for not giving a power-point presentation, saying, falsely, that the only time he had given one it was a complete disaster. Believing him, the audience 'competently deduced' and thus came to acquire the justified belief that he had never given a successful power-point presentation. That belief, while true, was made true by the fact that he has never given a power-point presentation at all. According to Williamson, the listeners,

basing their justified true belief that I had never given a successful power-point presentation on their justified false belief that the only time I had given a power-

---

asserting that that pumpus is a grumpus, you are saying merely that there is some grumpus in the vicinity. Some languages mark the difference by different words or different word order: German "Da ist ein..." v. "Es gibt ein...;" Hungarian "Ott van egy..." v. "Van ott egy." It is not the first time that the fly was lured into the fly-bottle by taking a quirk of English at face value. (John Biro, "What Is 'That'?" *Analysis* 71, 4 (2011): 651-653.) I am not denying, of course, that the proposition expressed by a sentence containing the first locution entails the one expressed by a sentence containing the second. I am claiming only that believing the first proposition does not entail believing the second, so that the latter's turning out to be true is not enough for the conditions of the traditional analysis to be satisfied.

<sup>18</sup> Williamson, *Philosophy*, 183.

point presentation it was a complete disaster... they did not know that I had never given a successful power-point presentation.<sup>19</sup>

I think that if we allow that the audience's false belief is justified, we have to say that it does know that Williamson had never given a *successful* power-point presentation, which is the only belief it acquired from his apology and which is both justified and true. (In VI. below I question whether we should allow this.) The fact that it is also true that he had never given one at all, something that entails what the audience believes but is not entailed by it and something that the audience does not believe, is neither here nor there. Of course they do not believe *that*, given Williamson's pretence.

Hercule Poirot to the assembled company: "Someone in this room is the murderer!" Poirot may or may not have a particular person in mind. Suppose he does. Are we to think of him as believing that if that person turns out to have an iron-clad alibi, someone else in the room is the murderer? That would be to see him as having reason to believe that it could not have been someone other than those present. He may well believe this, but, surely, not on the grounds that it follows from his belief that his prime suspect did the deed. Even though "The nephew, standing by the fireplace, is the culprit" entails "Someone in this room is the culprit," believing the former does not entail believing the latter in a non-pickwickian way. To be non-pickwickian, the latter belief would have to survive the demise of the former; and it requires different evidence – at once, more and less.<sup>20</sup>

To think that Poirot would make such an inference would be to commit both the mistake we found in Gettier's first example and that we found in his second. We would have to see him inferring a general proposition from one essentially involving an individual, going from  $Fa \wedge Ga$  to one of the form  $(Ex) Fx \supset Gx$ . We would also have to think that since he believes  $Fa$ , and since  $Fa$  entails  $Fa \vee Fb$  and  $Fa \vee Fc$  etc., he would believe the disjunctions (and therefore believe that the nephew or the wife did the deed and that the nephew or the secretary did and so on). However, though both inference forms are obviously valid, thinking that believing the premise suffices for believing the conclusion assumes that

---

<sup>19</sup> Williamson, *Philosophy*, 192.

<sup>20</sup> Suppose, by contrast, that Poirot does not have someone in particular in mind but is confident that one of the company is guilty and hopes that he or she will give him/herself away. Now it does not matter who the culprit turns out to be. As long as someone in the room *is* the murderer, Poirot has a justified (we have assumed) and true belief, and there is no reason to deny him knowledge. But in such a case, he is not inferring his true belief from a false – albeit justified – one, as does Smith in Gettier's first case.

(recognized) validity is enough for basing one belief on another, that truth-preservation is sufficient for belief-preservation. That it is sufficient for the preservation of justification – the so-called Principle of Deducibility for Justification – has been denied (Thalberg), though the prevailing view seems still to be that it is. What I am calling into question here is an analogous principle for serious, non-pickwickian, belief. Such belief does not simply track logical relations.<sup>21</sup> While it is certainly constrained by them (for rational believers), it is also subject to other conditions. Thus neither existential generalization nor addition, valid forms of inference though they be, is enough to generate such belief. But since the belief condition of the traditional analysis is satisfied only if such a belief is present, Gettier's examples and those similar to it do not pose a threat to it.<sup>22</sup>

What about the beliefs involved in what Sorensen calls junk knowledge?<sup>23</sup> Can one not believe, and be justified in believing, that  $p$  or  $q$  only because one believes that  $p$ , no longer believing the disjunction once one ceases to believe the first disjunct? According to Sorensen, Smith knows the disjunction in Gettier's second example if only the first disjunct is true but does not know it if only the second is. What he has in the former case is "not a useful type of knowledge but... nevertheless knowledge."<sup>24</sup> But it is not the usefulness of what Smith has that is in

---

<sup>21</sup> Heathcote makes a parallel observation about evidence: "The idea that warrant transfers from one belief to another is to be in the grip of a false analogy: the analogy between evidence for a proposition and the truth of a proposition. Logical implication preserves truth but it does not preserve evidence for, and if one tries to force the analogy then evidence for, and hence the notion of justification, will end up being an epistemic concept that has truth-like properties... This has done untold damage in the history of philosophy, creating bad doctrine along with un-meetable demands." (Adrian Heathcote, "Truthmaking, Evidence of, and Impossibility Proofs," *Acta Analytica* (2014): 373) However, he does not recognize that the Gettier 'problem' is a case in point.

<sup>22</sup> Heathcote also says (surprisingly, in view of the passage just quoted) that to deny that Smith is justified in believing (h) requires "deny[ing] some tried-and-tested rules of logical inference." ("Truthmaking and the Gettier Problem," 164) This is to commit the very mistake he laments. We need not challenge the validity of the inference from  $p$  to  $p \vee q$  to deny that it is sufficient to yield serious belief. Distinguish between " $p$  or  $q$  or possibly both" and " $p$  or else  $q$ ." The former is true if either  $p$  is true or  $q$  is true or both  $p$  and  $q$  are true, whereas the latter is true only if just one of  $p$  and  $q$  is true and not true if both  $p$  and  $q$  are true. In ordinary discourse, " $p$  or  $q$ " is always understood in the second way, unless the speaker adds the third disjunct. Doing so is always odd ("She spoke in German or she spoke in French (or she spoke in both)") and sometimes not even possible ("She is in Paris or she is in Rome (or she is in both)").

<sup>23</sup> Roy Sorensen, "Dogmatism, Junk Knowledge, and Conditionals," *Philosophical Quarterly* 38 (1988): 433-54.

<sup>24</sup> Sorensen, "Dogmatism," 446.

question, but whether it is genuine knowledge at all. What more is he supposed to know in knowing the disjunction other than what he knows in knowing the first disjunct when that disjunct is true? If the answer is, as I suggest, nothing, the traditional analysis suffices for telling what Smith does and does not know. It is no surprise that junk belief would be enough for junk knowledge but not for the real thing. Just as calling something junk food conveys that it is not the genuine article, so calling something junk knowledge tells us that it is not real knowledge. It is only the latter, however, that the traditional analysis claims to capture.

IV. What, though, to make of the fake-barn and the stopped-clock cases? I have already noted (fn.4) that there is disagreement about whether these fit the Gettier mold. Take fake barns first. While it has been claimed that “it is almost universally accepted that the agent in the barn façade case lacks knowledge,” I am not alone in thinking that it is far from obvious that Henry, looking at a real barn in fake-barn country, does not know that he is seeing a barn, a view shared by at least Lycan and Hetherington.<sup>25</sup> After all, he has formed his belief that he is doing so in a non-defective way and without relying on any false assumption, explicitly or tacitly (as Gettier's Smith or the believer in the Lehrer's Nogot/Havit case have been alleged to have done.)<sup>26</sup> But wait: is he not assuming that he is not in fake-barn country? Yes, but that is different from assuming that Nogot owns a Ford on the way to concluding that someone in the office does. The latter assumption is a positive contributor to a belief that would not be formed without it. The assumption that things are normal, in the absence of any reason to think that they are not, plays no such role in the fake-barn case. It is merely an instance of a general non-skeptical assumption, just like the assumption that there is no evil demon or that one is not a brain in a vat.

Pritchard says about the fake-barn case that “...there are... a great many nearby possible worlds where Henry forms the same belief on the same basis (by simply looking at the ‘barns’) and yet his belief is false.”<sup>27</sup> The worlds Pritchard has in mind are ones in which Henry is in fake-barn country, does not know it, and comes to believe that he is seeing a barn on the basis of a glance at a mere façade. But why think these worlds relevant? In them, one of the three conditions of the traditional analysis is not satisfied. Why should that be thought to impugn

---

<sup>25</sup> William G. Lycan, “Evidence One Does Not Possess,” *Australasian Journal of Philosophy* 55 (1977): 114-126; Hetherington, “Knowing Failably.”

<sup>26</sup> For a catalogue of the ways in which these cases may be modified in such a way that the no-false-assumption condition is satisfied, see Lycan, “On the Gettier.”

<sup>27</sup> Pritchard, *Epistemic Luck*, 162

the sufficiency of the analysis when the conditions *are* satisfied? Why should the fact that Henry might have been deceived deprive him of knowledge when he is not? We are allowed to assume that there is no evil demon. Still, there might have been. Does that force us to be skeptics?

The supposition that Henry is unaware that he is in fake-barn country is crucial here. In supposing Henry to be ignorant of being in a world as remote from the actual world as is a fake-barn world, we render his being in that world irrelevant to whether he has knowledge. Should Henry have some reason to suspect that he may be in fake-barn country, things would be different.<sup>28</sup> Then he would fail to satisfy the justification condition if he formed his belief merely by glancing at a facade. Nor would he, if sensible, believe his eyes. (Here we have the same trade-off we saw with Smith in Gettier's first case. See fns. 7 and 9 above.)

This is where the stopped-clock case is different from the fake-barn case in which Henry has no reason to be suspicious. In the stopped-clock case there is a specific tacit assumption in play, not just a general non-skeptical one. We cannot be supposed to be ignorant of the fact that in the actual world there are inaccurate or non-working clocks, hence anyone who forms a belief, even a true one, on the basis of a mere glance, fails the justification condition. That the clock I have just glanced at is a working and accurate one is a far more vulnerable assumption than a no-evil-demon one, and a world in which it is not is far closer to the actual world than is a fake-barn one. We all know that clocks are often inaccurate and sometimes stop, and we know how to go about finding out whether they are or have. (How does one go about finding out whether an evil demon is at work?)

Thus with clocks, whether the justification condition is satisfied depends on whether one has taken sufficient care to make sure that one is looking at a working and accurate clock.<sup>29</sup> I would not think I had done so just by glancing at a

---

<sup>28</sup> We are in the days of the land rush. Put something on a section and you own it.

<sup>29</sup> What counts as sufficient care is arguably context-dependent. See Igor Douven, "A Contextualist Solution to the Gettier Problem," *Grazer Philosophische Studien* 69, 1 (2005): 207-228 and Biro, "Showing." On the other hand, it may be held that there is a matter of fact about whether someone is sufficiently justified to know that is independent of his (or his attributor's) context, the latter affecting only how much he does or should care about whether he does satisfy it. Much of the time it does not matter whether one acts on knowledge or reasonable belief.

Harman's assassination case admits of being treated along similar lines. It is assumed that Jill, unaware of the false retractions of a true report of an assassination, satisfies the justification condition, though she would have believed the retractions, and thus ceased to believe the original reports, had she seen them, as did others who did. But that depends. (As usual, the case is severely under-described.) True, "Don't believe everything you read in the papers!" is not based on worries about evil demons – remember "Dewey defeats Truman"! But

clock, even one that has worked and been found accurate in the past. I may well form a belief about what the time is on the basis of a casual glance, but would I claim to *know* it, especially, did something important – say, your detonating the charge at just the right moment – hang on my doing so? Indeed, would I claim to *seriously* believe it?

It has not been noticed, to my knowledge, that in one way the fake-barn case and the stopped-clock case are inverses of one another. In the former, beliefs based on glances at all but one of the barn façades are false. In the latter, beliefs based on all but one glance at the clock – that is, all the past ones, at various times, at the (let us suppose) working and accurate clock – are, other things being equal, true. (I say ‘other things being equal,’ to allow for mis-readings and the like.) This is to be distinguished, of course, from the fact, often noted, that glances at the stopped clock at other times would have yielded false beliefs.<sup>30</sup> We can then imagine a clock case that mirrors the fake-barn case, one in which my every past glance yielded a false belief, since each was, as it happens, a glance at an inaccurate (or non-working) clock. But the clock has been repaired and when I glance at it again it shows the right time. Now suppose that I am justified (by whatever standard is appropriate in the various contexts) in forming the beliefs I do, on both those past occasions and on the present one. This can be so, even if it often requires more than a casual glance. My past beliefs were false, hence I did not satisfy the truth-condition of the traditional analysis. But on the present occasion I do satisfy all three conditions. There is, I claim, no more reason to say that I do not know what time it is than there is to say that Henry does not know that what he is looking at is a barn. Should I be aware, however, of the past unreliability of the clock, the bar for meeting the justification condition would rise dramatically, just as it would for Henry if he had reason to suspect that he is in fake-barn country. Neither situation is one in which it is clear that I satisfy all three conditions laid down by the traditional analysis yet lack knowledge.

V. Let us now consider the claim that that not only is there a genuine Gettier problem but that it is, in Zagzebski's word, inescapable.<sup>31</sup> She argues that the traditional analysis, even if supplemented by a condition or conditions of the sort

---

what if Jill lives in North Korea? Seeing the later reports with their stock footage of the dear leader would, far from leading one to doubt the original reports, re-inforce one's confidence in their veracity. (“They must have slipped one by the censors!”)

<sup>30</sup> E.g., Pritchard, *Epistemic Luck*, 156.

<sup>31</sup> Linda Zagzebski, “The Inescapability of Gettier Problems,” *The Philosophical Quarterly* 44, 174 (1994): 65-73.

often proposed to block Gettier-style counterexamples, would still be open to these. She offers what is, in effect, a recipe for constructing Gettier cases:

...start with a case of justified (or warranted) false belief. Make the element of justification (warrant) strong enough for knowledge, but make the belief false... due to some element of luck. Now emend the case by adding another element of luck, only this time an element which makes the belief true after all. We now have a case in which the belief is justified (warranted) in a sense strong enough for knowledge, the belief is true, but is not knowledge.<sup>32</sup>

Zagzebski gives us a number of examples in support of this sweeping claim. Here is one, aimed at Plantinga's attempt to solve the Gettier problem by appending to the traditional analysis a condition requiring that the subject's faculties be working properly in an appropriate environment. Imagine that Mary has very good eyesight – good enough for her to see, and thus to come to know, when her husband is sitting in this favourite chair in the living room. But on this occasion, Mary's husband's brother, who looks a lot like the husband, is sitting in the chair. Mary is, of course, entitled to conclude from her false belief that her husband is sitting in his favourite chair that he is in the living room. As luck would have it, he is. Thus Mary's belief that her husband is in the living room is true, but, intuitively, it is not an instance of knowledge.

It should be clear that what must be said about this example is just what I argued must be said about Chisholm's. Mary's belief concerns the person she is seeing, not one she is unaware of. That she thinks that that person is her husband is just like thinking that the bush one is looking at is a sheep. If the example exhibits the general form of Zagzebski's recipe for Gettier-style counter-examples even to a strengthened traditional analysis, a reply along these lines will always be available. This also shows, of course, that such strengthening is not required.<sup>33</sup>

Another example Zagzebski gives is that of Dr. Jones, who

...has very good inductive evidence that her patient, Smith, is suffering from virus X. Smith exhibits all of the symptoms of this virus, and a blood test has shown that his antibody levels against virus X are extremely high. In addition, let us suppose that the symptoms are not compatible with any other known virus, all

---

<sup>32</sup> Zagzebski, "Inescapability," 69.

<sup>33</sup> It should be obvious that the example fails the 'namely' test. The proposition that the husband is sitting in his favourite chair entails that he is in the living room. It also entails that he is somewhere in the house. But while his being in the attic makes the latter proposition true, surely, that is not what Mary believes. Suggest to her that what she believes is that her husband is somewhere or other in the house, and see what she says. Explain to her that she is wrong because this follows from what *she* says she believes, namely, that her husband is sitting in his favourite chair, and she will say, "You philosophers!"

of the evidence upon which Jones bases her diagnosis is true, and there is no evidence accessible to her which counts significantly against the conclusion. The proposition that Smith is suffering from virus X really is extremely probable on the evidence. None the less, let us suppose that the belief is false. Smith's symptoms are caused by the presence of a different and unknown virus; the antibody levels are due to idiosyncratic features of his biochemistry which cause him to maintain unusually high antibody levels long after a past infection. In this case Dr. Jones' belief that Smith is presently suffering from virus X is false. ... Now to construct a Gettier-style example we simply add the feature that Smith has very recently contracted virus X, but so recently that he does not yet exhibit symptoms caused by X, nor has there been time for a change in the antibody levels due to this recent infection.<sup>34</sup>

The trouble here is that the belief Dr. Jones supposedly forms on the basis of Smith's symptoms is under-described by Zagzebski. It is not just that Smith has X but that X is what is causing the symptoms. This is not a belief that is, or could be, made true by the fact that a-symptomatic X is present. Obviously, Dr. Jones would never form the belief that Smith has a-symptomatic X on the basis of his evidence. Thus the fact that Smith does have a-symptomatic X is not relevant to the truth of the belief Dr. Jones does form, a belief that is, by hypothesis, false. Once again, it is only by (perhaps inadvertent) slight of hand that we are maneuvered into agreeing that Dr. Jones has a justified true belief. But if he does not, we do not have a Gettier case.

I have not offered a direct argument against Zagzebski's general claim that no fallibilist account of knowledge can escape the Gettier problem.<sup>35</sup> But if my misgivings about her pivotal examples is well founded, we have reason to be skeptical about it.

VI. There remains one more putative Gettier case to comment on: Turri's amended Lamborghini story. In its initial version it goes like this:

One of Dr. Lamb's students, Linus, tells her that he owns a Lamborghini. Linus has the title in hand. Dr. Lamb saw Linus arrive on campus in the Lamborghini each day this week. Linus even gave Dr. Lamb the keys and let her take it for a drive. Dr. Lamb believes that Linus owns a Lamborghini, and as a result concludes, "At least one of my students owns a Lamborghini." As it turns out, Linus doesn't own a Lamborghini. He's borrowing it from his cousin, who happens to have the same name and birthday. Dr. Lamb has no evidence of any of this deception, though. And yet it's still true that at least one of her students

---

<sup>34</sup> Zagzebski, "Inescapability," 71.

<sup>35</sup> Heathcote ("Truthmaking, Evidence") takes a good stab at one.



owns a Lamborghini: a modest young woman who sits in the back row owns one. She doesn't like to boast, though, so she doesn't call attention to the fact that she owns a Lamborghini.<sup>36</sup>

As Turri notes, this story is 'vaguely modeled' after Lehrer's Havit/Nogot story, and, it should be easy to see that it succumbs to the same objection: it fails the 'namely' test. But Turri gives it a twist that clearly blunts that objection. He asks us to imagine that "...unbeknownst to Linus he has just inherited a Lamborghini. His cousin died and left it to him."<sup>37</sup> Here there is no questioning that Lamb believes exactly what happens to be true, namely, that *Linus* owns a Lamborghini. However, there is another problem, both with the original and the amended version. In neither is Lamb's belief that Linus owns a Lamborghini justified, at least not in the sense needed for satisfying the traditional account. The fact that Lamb does not realize that she is being deceived may explain and excuse her belief. But that is not the same thing as providing adequate grounds for holding it, which is what, I submit, the traditional account requires. Swallowing lies, even if blamelessly, is not a good reason for believing, and a belief so arrived at is not a justified one. For a belief to be justified in the relevant sense, it must be the case that the evidence the believer has supports the it, where what that means is that the evidence's being what it is makes the proposition believed more likely to be true than it would be otherwise. Someone may be forgiven for believing what he does, even if the evidence on which he bases his belief does not support that belief, as long as it appears to support it and the believer has taken sufficient care in assessing it.<sup>38</sup> If we assume that there is no evidence available to Lamb that would suggest that Linus may be deceiving him, he may be seen as blameless. Nonetheless, Linus' lies are not evidence that he owns a Lamborghini. If so, the justification condition of the traditional analysis is not satisfied.<sup>39</sup>

In fact, deliberate deception is not the only thing that can rob a believer of the kind of justification the traditional account says is needed for knowledge. A simple mistake can do so. Imagine that Lamb has the same evidence of Linus' ownership except for Linus' telling her that the car in question is a Lamborghini.

---

<sup>36</sup> Turri, "In Gettier's," 215.

<sup>37</sup> Turri, "In Gettier's," 271.

<sup>38</sup> Whether it does depends on whether he has taken as much trouble to make sure that he is not being deceived as the situation demands, which, in turn, depends on what is reasonable for him to believe about his students in general, Linus in particular, and so on.

<sup>39</sup> Given Iago's machinations, it is understandable that Othello takes Cassio's possession of Desdemona's handkerchief as evidence of her infidelity. It is because it is not that we, who know that the 'evidence' is planted, say that even if she were in fact unfaithful, he would not know it.

It is, in fact a Maserati, but one has to be quite knowledgeable about Italian cars not to mistake one make for the other. Lamb is not and comes to believe that Linus owns a Lamborghini. As it happens, Linus does own both the Maserati he shows Lamb and a Lamborghini he keeps in his garage. Gettierites will be quick to say that Lamb does not know that Linus owns a Lamborghini, and I agree. But the reason why he does not is not, as they claim, that having a justified true belief is not sufficient for knowing. It is that he is clearly not justified in believing what he does.

A similar twist may be given to Gettier's own first case. Suppose Smith, having been asked by Jones to hold the latter's jacket while he is changing a flat tire, is furtively checking all the pockets to see how much money Jones has. He thinks he has counted ten dimes and thus forms the belief that Jones has a dollar. However, in his hurry he missed one of the pockets but counted one of the dimes in another twice. Even if the missed pocket has a dime in it, Smith is not justified in believing that Jones has a dollar, and *that* is the reason he does not know that the man who will get the job has a dollar in his pocket, even if that man turns out to be Jones.

What happens in such cases is that the justification condition is not really satisfied, even though it appears to be to the subject, because the evidence the subject is relying on is not probative of the proposition he believes. To be probative, the evidence must in fact support the proposition, not just be taken by the believer to do so.<sup>40</sup> The traditional analysis is not threatened by such cases; it requires that the justification condition be in fact satisfied, not merely that it appear to be satisfied to the believer. (If evil-demon worries remain, they have nothing to do with the problem Gettier is alleged to have posed.)

In discussing Turri's example, I have challenged his claim that in it the justification condition is satisfied (as I suggested it may be the case with the stopped clock). That is a different complaint from my earlier one that in the typical Gettier examples it is the belief condition that is not satisfied. But, as I noted before, the real question is whether in any of these examples *both* conditions are, as they need to be if the cases are to be counterexamples to the traditional analysis.

---

<sup>40</sup> To say that the belief must in fact be justified is not to abandon fallibilism about justification, a desperate measure advocated by some (Robert Almeder, "The Invalidity of Gettier-Type Counterexamples," *Philosophia* 13, 1-2 (1983): 67-74; Scott Sturgeon, "The Gettier Problem," *Analysis* 53, 3 (1993): 156-164; Trenton Merricks, "Warrant Entails Truth," *Philosophy and Phenomenological Research* 55, 4 (1997): 841-855.) As stressed by Lewis, doing so quickly leads to skepticism. (David Lewis, "Elusive Knowledge," *Australasian Journal of Philosophy* 74, 4 (1996): 549-567.) Being probative should not be confused with being factive.

Even if one's justified but mistaken belief entails a true proposition, and one recognizes that it does, it does not follow that one must come to believe the entailed proposition. One way to see this is through seeing what is wrong with the following, seemingly plausible, objection: Surely, someone can come to believe  $Q$  if she believes  $P$  and believes that  $P$  entails  $Q$ . Why not? Can we not infer new beliefs from old beliefs, recognizing logical connections? Is that not what is called deductive reasoning? Why is it not possible for someone to infer from  $Fa$  to  $ExFx$ , and to really believe, and not just 'pickwickian believe,' that  $ExFx$ ?

Of course, it is, in most cases. But not if the inferred belief can be true even if the belief from which it is inferred is false. Suppose Poirot says "someone in this room is the murderer" because he believes that the nephew killed the uncle. On subsequently discovering that the nephew has a cast-iron alibi and it was the butler, also present, who committed the dastardly deed, we would not allow Poirot to get away with saying (not that he would), "I was right all along!" Similarly, we would think it a poor joke if Chisholm's sheep-spotter claimed that he had a true belief all along.

There is no harm in saying that the inferred proposition is believed when the proposition from which it is inferred is false as long as we remember that it is believed only in a manner of speaking. But the acceptability of such a *façon de parler* should not be allowed to mislead us into thinking that the belief condition of the traditional analysis is satisfied with respect to the proposition that turns out to be true.

Something like this is recognized by Gaultier in a recent paper in which he argues that it is impossible for Smith to form the belief that is made true by his getting the job and having ten coins in his pocket:

... when the belief that John owns a Ford has been formed in the way indicated in the description of the Gettier case, this belief cannot lead one to form, in addition, the different belief that someone in the company owns a Ford.<sup>41</sup>

The reason, according to Gaultier, is that in general

...one cannot believe about the question whether  $p$  something weaker, more indefinite or undetermined, than what... [one's evidence] appears to one to support or establish about the question whether  $p$ .<sup>42</sup>

<sup>41</sup> Gaultier, "An Argument," 267.

<sup>42</sup> Gaultier, "An Argument," 267. (While Gaultier's discussion is confined to Gettier's first example, it is easy to see how it can be extended to the second: " $p$  or  $q$ " clearly expresses a weaker claim than does " $p$ ." Gaultier thinks that a commitment to the voluntariness of belief is essential to his argument. I am not sure that I see why. But it is not essential to mine. He also

I think Gaultier is on to something; however, through failing to distinguish between serious and pickwickian beliefs, he commits himself to a stronger thesis than is needed to block the supposed counterexample. Consider again the sheep-in-the-field case. While the proposition expressed by “*There* is a sheep” entails the proposition expressed by “There is a sheep (somewhere),” believing the former does not entail seriously believing the latter. If someone who believed the former were asked whether he believed the latter, no doubt he would say that he did – though in a special tone of voice, I think. We can allow that there is a sense – the philosopher’s sense – in which one believes any proposition one thinks is entailed by a proposition one believes. But we need not agree that one seriously believes everything one thinks is entailed by something one believes.

Thus if by ‘belief’ we mean ‘serious belief,’ Gaultier is right that one cannot infer a ‘There’ belief from a ‘*There*’ belief. But this is not just, or primarily, because the former is weaker. He offers as a general principle that

...at *t*, one cannot believe about the question whether *p* something weaker, more indefinite or undetermined, than what, at *t*, *E* appears to one to support or establish about the question whether *p*.<sup>43</sup>

This seems to me to be too strong a claim. Having just seen my neighbour enter his house, I believe that there is someone in that house; I will bet you that there is if you claim otherwise. In fact, my neighbour has left through the back door but a burglar has snuck in the same way. The reason why this is not a Gettier case is not that believing that there is someone in the house is weaker than believing that my neighbour is in the house – though it is – but because believing that there is someone (i.e., my neighbour) in the house is not the same thing as believing that there is someone or other in the house.

A mark of seriousness in a belief is that it guides action. If I am in the market for a used Ford and believe that Havit owns the one in the parking lot, it would not be rational for me to go around asking who owns it, as it would be if what I seriously believed was that someone or other in the building did. And if I believed that Secretariat was a dead certainty to win the 1973 Belmont, even though I realized that from the proposition that Secretariat will win it follows that either Secretariat or Twice a Prince, the rank outsider, will win, it would not be rational to split my bet between the two. Not only that – taken as serious beliefs,

---

thinks, as I do not, that perhaps the fake-barn case and certainly the stopped-clock case are counterexamples to the traditional analysis.

<sup>43</sup> Gaultier, “An Argument,” 267.

the two in each pair are incompatible, as shown by the fact that I cannot act on both.<sup>44</sup>

VII. To have a counter-example to the traditional analysis, we need a case where it is clear that one satisfies the conditions it lays down as sufficient for knowledge and it is also clear that one lacks knowledge. I have argued that Gettier's original cases and those that follow the same pattern do not succeed as counter-examples because in those cases one of the three conditions is not satisfied. In many, the subject does not seriously hold the belief that turns out to be true. With some (fake-barn, stopped-clock), I am content with a Scotch verdict: it is at least not *clear* that they are counter-examples to the traditional analysis, because it is not clear that the subject lacks knowledge.

I have no proof that a case cannot be described (not even that one has not been) that does not exhibit one of these three patterns and in which it is clear that all three conditions are satisfied and also that the subject does not know. But I do think that seeing the ways in which the well-known cases I have discussed fail strongly suggests that the confident claims that the traditional analysis has been decisively refuted should be treated with caution. There is, perhaps, life in the old analysis yet.

---

<sup>44</sup> It gets worse. If I believed, on the basis of recognizing that it follows from my belief that Secretariat would win, that either Secretariat or Twice a Prince would win, I would have to believe that Secretariat or Sham would win and that Secretariat or My Gallant would win and that Secretariat or Private Smiles would win. (There were only five runners.) Ten dollars on Secretariat would have netted me a profit of two dollars; splitting my ten dollars five ways, a loss of seven dollars and eighty cents. In the 1978 Belmont, even though I believed, along with many, that, given the extra furlong, Alydar would finally catch up with Affirmed, I did also believe that either Alydar or Affirmed would win. But that was not by way of deducing the second proposition from the first. It was because there was independent reason to believe the latter, as reflected in the respective odds (6/5 and 3/5). I did not also believe that either Alydar or Darby Creek Road or that either Alydar or Judge Advocate or that either Alydar or Noon Time Spender would win. In fact, believing any of these is incompatible with believing what I did, as is believing any one of Gettier's 'constructed' disjunctions is with believing one of the others.



# TRACING THE TERRITORY: A UNITARY FOUNDATIONALIST ACCOUNT

Olga RAMÍREZ CALLE

**ABSTRACT:** The paper offers an integrative interpretation of the different lines of thought Wittgenstein was inspecting in *On Certainty* and what he might have been looking for through them. It suggests that we may have been focusing our attention too strongly in the wrong place and comes to a new conclusion about where the real import of these reflections lies. This leads to an answer to the initially posed question of foundationalism that revises the way in which there can be said to be a grounding intention in *On Certainty*.

**KEYWORDS:** foundationalism, Ludwig Wittgenstein, *On Certainty*, hinges

## 1. Introduction

Ours is an inherited world; beyond the purely physical there are so many sediments of understanding that we cannot even imagine making any sense of our surroundings without some guidance. Not just knowledge and experience about nature and how to deal with it, but our own self-understanding and the way to behave with and towards others in society, are due to history. This was the deep insight brought about by late phenomenology. Even if we reform, revise and renew, there is so much we take for granted that we can't even begin to become aware of it. It is not just that we assume a world with such saturated drawn contours; it is that we would not even come close to the idea of questioning much of it, at least not regularly.

We rely on steadfast linguistic meaning to speak about our environment, develop theories, establish connections between ideas, calculate and measure. This built-in vocabulary that we take for granted, already records deposits of information, condensed hypotheses, theories (and errors), pragmatically guided distinctions, customary evaluations, and much more. On the basis of the meaning introduced through explicit or implicit definitions of our terms we consider some claims as analytic. We regard them as true in virtue of the meaning of their constitutive terms. Ordinarily we don't question the linguistic tools we use; the

truth of our statements relies on the stability of meaning, of linguistic rules – which include, of course, logical ones. We could not make assertions, nor gain further knowledge, without them. We must keep them fast in order to assess, interpret, and gather new information about the world.

Wittgenstein's *Philosophical Investigations* entrench with this stream of thought but he is said to have carried it further in *On Certainty*;<sup>1</sup> anchoring this picture of the evolving flow of human gnoseological and linguistic patrimony, or so it is argued, to some new kind of foundation 'with a human face.' This would be less than epistemological or logical in classical terms, but would rather draw the boundaries of senselessness for human beings. It is my aim in this paper to try to see to what extent the considerations Wittgenstein brings up in *On Certainty* advance a specifically new phenomenon in the foundational sense previously referred to.

## 2. Hinges and the Question of Foundationalism

401. I want to say: propositions of the form of empirical propositions, and not only propositions of logic, form the foundation of all operating with thoughts (with language) (...). (Wittgenstein, *On Certainty*)

The contemporary discussion around Wittgenstein's *On Certainty* has majorly focused on what Wittgenstein called 'hinges.' Hinges are said to be certainties in the sense that they are assumed to be true, or, perhaps, cannot but be assumed to be true, rather than being a priori known to be so or being epistemically warranted. Although they are empirical statements they are not in the market for justification or rebuttal. Even if they belong to the normative background of our living they differ from the a priori statements, such as statements of logic and what Wittgenstein called 'grammatical statements' in that they are contingent.

The standard statements Wittgenstein provides as examples in the context of *On Certainty*, statements such as "The Earth has existed for many years past," "Here is a hand," "My name is L. Wittgenstein," "I just had lunch," "There is a staircase outside my bedroom," "I have never been to Asia Minor," "No one has ever been to the Moon,"<sup>2</sup> or "Objects don't disappear when they are not looked at" etc. would be undoubtedly accepted for reasons other than proof or any clear

---

<sup>1</sup> Ludwig Wittgenstein. *On Certainty*, eds. Elisabeth Margaret Anscombe and George Herbert Von Wright (New York: Harper Torchbooks, 1972. First published 1969)

<sup>2</sup> Said, of course, in Wittgenstein's time. A contemporary equivalent could be: "No one has ever been to the Galaxy Cassiopeia Dwarf."



epistemic justified necessity in their support. It is rather, so it is argued, that we could not make sense of our lived world without them being true.

The question of whether there were foundationalist intentions in Wittgenstein's last work has also been at the centre of the recent debates. Positions have ranged from those that do ascribe hinges a foundational character<sup>3</sup> to those that rather see Wittgenstein as finally reducing his aspirations to signaling some heterogeneous constitutive, taken for granted statements in our different life contexts that are different from mere grammatical or logical ones and whose acceptance is not in question. Those that do not defend a foundational reading tend to see at best some form of coherentism in Wittgenstein's remarks or deny the apprehension of his position in any of these categories.

The difficulties posed to an homogeneous foundationalist reading find support in some puzzling aspects of the text itself: the different character of the propositions offered as examples of basic hinges (which range from clearly more basic ones to others that would not so easily be considered universally shared), the differing levels of research this implies and the diverse perspectives from which they are approached. The propositions offered as examples of hinges include:

- a. The most basic assumptions about the physical world, without which we could not live or move around without continuous disconcert, accidents and suicidal experiences include: "Objects don't disappear when I am not looking at them," "There are physical objects," "I am here right now," "Human beings have parents."
- b. The most basic assumptions about the correctness of our logical and mathematical operations with numbers: " $12 \times 12 = 144$ ."
- c. Recognitional statements,<sup>4</sup> statements such as Moore's: "I have two hands," "There is a tree" etc.

---

<sup>3</sup> Among the supporters of the foundationalist reading we find for example, Danièle Moyal-Scharrock, "Unraveling Certainty," in *Readings of Wittgenstein's On Certainty*, eds. Danièle Moyal-Scharrock and William H. Brenner (New York: Palgrave Macmillan, 2005), 76-99, and also, at least to some extent, Crispin Wright, "Skepticism, Certainty, Moore and Wittgenstein" in *Wittgenstein's Lasting Significance*, eds. Max Kölbel and Bernhard Weiss (London: Routledge, 2004), 228-248. Among those that reject a foundationalist reading we can count Michael Williams, "Why Wittgenstein isn't a Foundationalist", in *Readings of Wittgenstein's On Certainty*, 47-58 and also Annalisa Coliva, "Hinges and Certainty. A Précis of Moore and Wittgenstein. Skepticism, Certainty and Common Sense" in *Philosophia*, 41 (2013): 1-12.

<sup>4</sup> I have adopted the expression 'recognitional statements' from Williams, "Why Wittgenstein," 49, since it seems to me to fit well and clarify the special character of this kind of statements.

- d. Statements about personal life and identity: “My name is L.W,” “I come from town x,” “there are stairs outside my bedroom,” “I have never been to Asia Minor.”
- e. Epistemological boundary statements: “No one has ever been to the moon.”
- f. Statements about our gnoseological patrimony: comments on chemistry, physics, history, textbooks in general, etc.
- g. Statements about general background assumptions in all kind of contexts and possible language games.

We may have the tendency to select just those examples of statements that most directly favor what we might call a ‘basic reading,’ supposing that the rest is less essential to the point. But before we do, we should first consider whether we could integrate the whole as research lines pertaining to a common project. If we adopt this strategy the truth is that sometimes Wittgenstein’s remarks on those ‘contingent statements that would adopt a normative rather than an epistemic role,’ allow a much wider reading. Towards the end of the listing above we tend to depart from those very basic propositions whose questioning by any human being would tend to make his life appear absurd to less dramatic ones. The corresponding last remarks appear at times to refer to any contingent statements that fit the bill. So one might wonder, if they weren’t of interest for the inquiry in *On Certainty*, why include them, especially if he had already done so in PI?

We cannot simply ignore the fact that sometimes he does seem to be referring to no more than those general background assumptions, transmitted knowledge and layers of understanding I referred to in the introductory passages, many of which are ingrained in our (evolving and disparate) conceptual baggage, and are not universally shared and would not, as Williams argues,<sup>5</sup> satisfy other requirements of an adequate basic, foundational set of statements, which according to him would include: being distinguishable as basic, being universal, independent of the body of knowledge, grounding and allowing resolutions of disputes at a global scale. We could set apart statements from a) to e) as ‘the narrow reading of hinges’ (NRH) and include them all in what we might call in contrast ‘the broad reading of hinges’ (BRH). However, according to Williams not even as applied to NRH talking about foundationalism would be justified for the reasons mentioned.<sup>6</sup>

---

<sup>5</sup> See Williams, “Why Wittgenstein,” 50-58.

<sup>6</sup> I will not go into the reasons offered by Williams to defend this point, since he has done so extensively himself, *ibid*.

In the paragraphs where Wittgenstein embeds the question of hinges into his characteristic terminology of constitutive rules of language games, he clearly oscillates between examples that seem to embrace all kind of cultural and context-dependent games to others that suggest reference to the more basic hinges (NRH); that is, to shared assumptions of any human life whatsoever, which might differ in the specificity “I have never been to China” or “I have never been to the US,” “My name is LW,” or “My name is OR” but would basically count for the same.

Another turn is given with the passages that focus on what a proper response to the sceptic might be and why Moore’s answer isn’t an answer at all, sliding this way into what I take to be a relatively different problem stage.

All this makes it difficult to give an integrated interpretation of the text. However, in what follows I want to see to which extent even the less basic of these remarks are not just remnants of old thoughts but contribute to the introduction of a new epistemological perspective. One that might point to a unified foundationalist account, in which context most of these basic hinge statements (NRH) would be rather the most immediate fruitage than the ground itself.

### **3. Ideologies, Traditions, Theories and Presuppositions of Differing Sorts...Does Anything Count?**

If we depart from the most standard definition of hinges as those implicit normative but contingent propositions that stay put while others are used, enabling possible talking, discussing and inquiring into or about something else, we certainly face a very generalized phenomenon, since there are plenty of such contingent empirical statements we take for granted in our diverse life contexts. We find them from the most intimate micro-historical personal world levels, such as between me and my partner, me and my family, me and my friends, me and my colleague peers, to more professional contexts and larger communities (theories and ideologies of different sorts, psychoanalysis, physics or chemistry); to different community sizes and sorts; to shared socio-historical or biological ones. We share experiences and exchanges, transmitted knowledge about our life contexts which we take for granted, expecting others to do so too. Some of these might have a more restricted life than others but, nevertheless, as long as they are active, their role appears to be quite similar to that described in Wittgenstein’s examples. Consider his remarks about personal history: people knowing their names, their ancestors, where they come from, the spatial and time contexts in which they are living, where they have been, where they are located right now and, why not?,

whether they just met someone new, brushed their teeth, or other things of that sort.

159. (...) I believe that I had great-grandparents, that the people who gave themselves out as my parents really were my parents, etc. This belief may never have been expressed; even the thought that it was so, never thought.
431. "I know that this room is on the second floor, that behind the door a short landing leads to the stairs, and so on." One could imagine cases where I should come out with this, but they would be extremely rare. But on the other hand, I show this knowledge day in, day out by my actions and also in what I say.
70. For months I have lived at address A, I have read the name of the street and the number of the house countless times, have received countless letters here and given countless people the address. If I am wrong about it, the mistake is hardly less than if I were (wrongly) to believe I was writing Chinese and not German.
71. If my friend were to imagine one day that he had been living for a long time in such and such a place, etc. etc., I should not call this a mistake, but rather a mental disturbance, perhaps a transient one.

These remarks don't seem in themselves very different from similar personal expectations concerning my partner's knowledge that I have been working in this or that University, that I am a philosopher, etc., etc.; or that in Spain we drive on the right and not on the left, or that street signs ought to be respected; or, if my colleague is a psychoanalyst, that he doesn't suddenly ignore that for psychoanalysts there is something like the unconscious, or, as a life-long communist, that he ignores communism's main doctrines, and so on. Or are they? Some of them might concern information more essential to our living, but are the reasons for our reliance upon them (and our expectations towards others in that regard) any different? Am I less disturbed (or are those who share my life) if I cease to know that I am married and I am a philosopher and I work in this or that place than if I don't know my own address? Indeed it might not be that bad if I forget which university I graduated from, or the names of those that did so with me, but is it just a matter of importance and degree? Like when I take good care of the password for my bank account, which I need and keep reusing, while I forget others created in some unimportant, seldom visited websites, or remember the names of all my students while I deal with them, while some of them fade away with time... etc. Some of this is important for a lifetime, some just essential in short-term memory, for a while. Suspicion of Alzheimer arises when a family member starts forgetting the names of his sons, what his profession is, who his wife is, how many times he has been married, what he just did, where he was this

morning. When he forgets, that is, the course of his main activities throughout his day or life. But not when he doesn't recall whether he was ever in Albacete or not, unless he married and worked in Albacete at one time. Going to Asia Minor might count as a more exotic adventure to forget, but only if you are not some voyager travelling around the world every x number of days, who might once have landed, jet-lagged, in Asia Minor but cannot really say. So, what is the criterion?

Similarly, Wittgenstein devotes many passages to showing how much we rely on our more generally transmitted gnoseological patrimony and how it determines and constitutes all our further epistemological enterprises and dealings with the world. We do not start from zero in conducting experiments, we trust what anatomy and the history textbooks say.

167. It is clear that our empirical propositions do not all have the same status, since one can lay down such a proposition and turn it from an empirical proposition into a norm of description.

Think of chemical investigations. Lavoisier conducts experiments with substances in his laboratory and now he concludes that this or that takes place when there is burning. He does not say that it might happen otherwise, another time. He has a definite world-view – not of course one that he invented: he learned it as a child. I say world-view and not hypothesis, because it is the matter-of-course foundation for his research and as such unmentioned.

600. What kind of grounds have I for trusting text-books of experimental physics?

I have no grounds for not trusting them. And I trust them. I know how such books are produced – or rather, I believe I know. I have some evidence, but it does not go very far and is of a very scattered nature. I have heard, seen and read various things.

602. Should I say “I believe in physics,” or “I know that physics is true”?

603. I am taught that under such circumstances this happens. It has been discovered by conducting the experiment a few times. Not that that would prove anything to us, if it were not that this experience was surrounded by others, which combine with it to form a system. Thus, people did not make experiments just about falling bodies but also about air resistance, among other things. But in the end I rely on these experiences, or on the reports of them; I feel no scruples about ordering my own activities in accordance with them. – But hasn't this trust also proved itself? So far as I can judge – yes.

But not just such globally shared assumptions, actually them all, established theories and shared knowledge as well as more specific scientific paradigms and ideologies and other cultural idiosyncrasies, have mostly this cultural furnishing

role on whose stage we go on living, inquiring and forming further hypothesis. Some might be restricted to specific communities and cultures, others belong to the common patrimony of humankind. Among the latter we find those regarding the human biological species and the physical world, with statements of increasing informational complexity: from being mammals and learning grammar to vulnerability to atmospheric conditions, virus and bacteria of varied sorts; from our earthly venue to sophisticated theories about the universe as a whole.

However, they are all just steps in the ever narrowing circle of what each person in his personal sphere, in his community, his specialty, his culture, his species and his world takes for granted in living; from direct experiences, to mnemonically recorded ones and transmitted knowledge. Much built into, much sediment from which we depart in becoming active in the world; much belonging to some software deposits from which we draw in order to fulfill our active life but which do not directly belong to the forefront of our living. Those hinges relevant for the constitution of our identity and the world would seem to occupy the deeper, most permanent layers, in each of these stratified structures that we are.

#### **4. Grammar, Certainty, and Knowledge**

##### **4.1 Certainty of What?**

When talking about hinges beyond the importance that this taken for granted informational states have in our life, what is stressed is the certainty with which we rely upon them. We take hold of them with no hesitation, rely upon them almost blindly, how is this to be explained if not justified? Does the relevance they have for our life, or the more or less dependence of whole structures of knowledge upon them, makes them trustworthier? Is there some epistemic basis for our trust or is trust just a function of necessity? Is our certainty concerning the different sort of hinge-statements of the same sort? Or are local ones, for example, less certain in some sense, than universally hold ones?

The wide range Wittgenstein is ready to give his notion of hinges shows itself again in those passages where he more explicitly poses the problem in terms of constitutive rules of language games. Where he offers examples referring also to local, cultural, historical and possibly changing contexts. Actually, given the ample notion of language games he entertains this is simply a matter of consistency. Since we do find those newly acknowledged empirical statements that play the role of constitutive rules in most language games whatsoever and take them for granted just the same as long as we move within their frames.

Consider these remarks:

620. In particular circumstances one says “you can rely on this;” and this assurance may be justified or unjustified in everyday language, and it may also count as justified even when what was foretold does not occur. A *language-game exists* in which this assurance is employed.
609. Supposing we met people who did not regard that as a telling reason. Now, how do we imagine this? Instead of the physicist, they consult an oracle. (And for that we consider them primitive.) Is it wrong for them to consult an oracle and be guided by it? – If we call this ‘wrong’ aren't we using our language-game as a base from which to combat theirs?
617. Certain events would put me into a position in which I could not go on with the old language-game any further. In which I was torn away from the *sureness* of the game. Indeed, doesn't it seem obvious that the possibility of a language-game is conditioned by certain facts?

These reflections go on in §579, §628 and others.<sup>7</sup> In such cases it would not be surprising to hear that our reliance need not depend upon the truth of the relied upon. However, interpreters have stressed the relevance of realizing that the ‘sureness’ Wittgenstein is talking about in talking about hinges is altogether of a non-epistemic sort. It is not a matter of knowing the statements in question to be true. What a community assumes to be the case can turn out false from an epistemic perspective. It would be ignorance not to recognize the numerous examples by which taken for granted statements were discovered wrong and changed throughout history. This notwithstanding that whole cultures relied upon them as their unquestioned background. Nor is Wittgenstein a relativist in the epistemic sense as can be derived from §617. So, first of all, if there is something like being certain in these cases too, what is it that I am certain of? and further: Can I be said to have the same kind of certainty regarding some taken for granted beliefs in my local community, or in the world community at large, than I have for the kind of personal hinges we saw above? And, most importantly, if not, why not?

---

<sup>7</sup> Wittgenstein's *On Certainty*, §579. It is part of the language-game with people's names that everyone knows his name with the greatest certainty.

§ 628. When we say “Certain propositions must be excluded from doubt,” it sounds as if I ought to put these propositions – for example, that I am called L.W. – into a logic-book. For if it belongs to the description of a language-game, it belongs to logic. But that I am called L.W. does not belong to any such description. The language-game that operates with people's names can certainly exist even if I am mistaken about my name, – but it does presuppose that it is nonsensical to say that the majority of people are mistaken about their names.

Let's consider first the personal hinge cases again. We could draw the following distinctions between them:

- a) "I am here now," "I am sitting on a chair," "I am brushing my teeth," "I just saw a man," etc. and also: a)\* "There are stairs outside my bedroom," "I arrived today," "I had breakfast this morning" etc.
- b) "My name is LW," "I have been living at this address for long" etc.

The first a) allude ultimately to direct experiences I am having, which I simply cannot doubt, since all I do relies in my being capable to trust this kind of things. If I couldn't, I would lose completely my sense of security and would not know what to trust any more. It is of course not that I have an external experience of 'myself doing this and that' but I do have a conscious awareness of '(my)<sup>8</sup> doing this and that,' '(my) experiencing this and that' I need not express it, nor claim to know it, but I have it and base myself on it. There is an enormous amount of experiences (beyond those Wittgenstein mentions) that I might be having and could include here. All my memories are based on my capacity to trust these very basic experiences too. Statements of the sort a)\* are already memories of these more immediate ones. Statements of the sort b) are based also on our memory of experiences, but of experiences of such a persistent sort and so relevant to my identity that they become essential to my living. They are tied to something that is not an episodic memory or a mere punctual experience, such as is the constant consciousness of what I call myself (whose alter-ego becomes LW, OR or whatever for the others). As Kant says, rather than an experience it accompanies them all as a form of continuous awareness. Were I to doubt myself as a whole, then surely nothing goes. But, to be sure, to know that "my name is LW" is not the same as this self-awareness, it is, as I said, a persistently recorded experience intimately tied to it. This sort b) could go beyond the cases mentioned to include many of those aspects relevant to my identity considered before: that I am married, a philosopher etc. The ground of my certainty, and there is such a ground here, is not of a different kind. Being OR is more essential to me than being married, but having been in Asia Minor need not be.

However, my certainty of experiences of the first sort a) or even of the second b), not in the most radical sense alluded before of myself accompanying my experiences, but in the sense of the recorded memory of a personal history I recall as my own (and I could possibly exchange with someone else's personal history, for example Napoleon's) would be just the same if I were hallucinating. I would rely upon them exactly the same. Since as a human being it is no option for me to

---

<sup>8</sup> In the sense of an internal awareness of myself as subject and not as object of experience



rely in my direct experiences but a necessity. Is it possible that I am wrong about my thereupon based beliefs? Yes, it is. Although others would be needed to make me aware of that. What is not possible is 1) that I am not certain of having had those experiences if I have (not of their being reliable if shown otherwise) and 2) that I could live trusting them without difficulties. Unless ... the whole community is having the same hallucinations and they are not incompatible with our remaining experiences about the world! But, in the normal case, if I certainly would be having hallucinations and there is a mismatch between my experiences (or my recalled personal history supposedly based on them), and my certainties, with the world, I would continuously have trouble with others and my environment. Nevertheless, I *prima facie* cannot but trust. It is not that I act as if the conceivable were possible; it is that the 'merely conceivable for others' is certain for me and that is why I act accordingly. More than a possibility it is a reality, though I can be wrong.

But now let's go back to the initial question of whether my certainties about transmitted and relied upon 'knowledge' in my community (however large) are of the same kind as those certainties about my experiences and memories based directly upon them. The claim is that those personal certainties do not have an epistemic character either and I would say this is true in a sense. It is true in the sense that I can deliver no further justification for them beyond repeating them again. That is, if knowledge requires to be justified as true then I could not deliver that in cases where I am registering experiences which themselves can be given no further justification. However, I am certain that these experiential states are the way to prove how things are and if this counts as knowledge, I have knowledge. This need not be incompatible with Wittgenstein's claim that saying "I know" adds absolutely nothing to my stating that *p* as I will soon elaborate. So let's distinguish for now between justified knowledge and knowledge\* of this last sort. Actually, as Timothy Williamson<sup>9</sup> would put it, knowledge (knowledge\*) is already in this first step of registering an experience as *p*. But now, still, in the case of relied upon transmitted information, what is it that I am certain of? Of course, I do take such informational states for granted in my life not necessarily having or asking for justifications. So the sureness with which I rely upon them has nothing to do with me having any kind of proof of its truth. My sureness is not epistemic in the sense that, for different reasons now, I need not be able to give justification here either in the traditional sense of proving knowledge. But, with that out of the way, there are still two different questions here: is my sureness about 1) the content of the statement in question being true (whether I have proof of it or not)

---

<sup>9</sup> See Timothy Williamson, *Knowledge and its Limits* (Oxford: Oxford University Press, 2000).

or 2) about my experience of this having been transmitted to me as something I could rely on? Take the case of the textbook of chemistry: what am I sure of? This being the way I have read things to be or of the content of those read statement being true justified knowledge (even without having been given the needed, maybe sophisticated justification)? Can I really separate both? If I compare the case to my personal experiences and memories, the only thing I can be said to be certain of in a similar way is 2). I have experienced this through textbooks or by seeing how my teachers or other society members apply it or take it for granted in different ways. That is, if this should be seen as a similar kind of certainty as my personal ones, the only thing I can be certain of is my experience that it is treated as knowledge, that I myself treat it as if it were true. I can be as certain of this as I am of sitting in a chair right now. If my experience of being sitting in a chair or having been sitting on a chair is knowledge\*, I can be said to have knowledge\* of those experiences in terms of 2) too. My certainty is a form of knowledge\*. But, when talking about certainty here we cannot reduce it so easily to that, since here the experience is one that concerns some cognitive content.

Let 's separate a few ideas.

- (i) I am sure about " $p$ " being true therefore  $p$  is the case.
- (ii) I am sure about having been transmitted " $p$  is true."
- (iii) I am sure about having been transmitted that  $p$  (is the case)

I do not have the kind of epistemic certainty of i) we said. I can be said to have some certainty and knowledge\* for ii), but since ii) implies that what I have been transmitted is that  $p$ . If based on ii) I simply rely on what I have been told or raised to rely on, I will be treating " $p$ " as if it were true. So if my certainty is just for ii) and iii) I will be acting as if i) were true. This requires the mediating step of me accepting and internalizing, trusting, that what I have been implicitly or explicitly taught to be true, is true. For this trust, there is in principle no certainty. That is, nothing like certainty of the type of knowledge\*. I do act as if I would have justified knowledge of the transmitted content. Furthermore I act as if the world would be as such statements say. But all I have in terms of certainties comparable to experiential and mnemonic ones, are of the type ii) and iii). Here as opposed to the personal experience cases, I could find out that I/we are wrong precisely by relying on those experiences of the first sort. In principle, though, I have no problem, as Wittgenstein says, "ordering my own activities in accordance to them" (to that transmitted knowledge) taking them to be as I am told, what amounts to taking them to be true. So, ultimately, there is something epistemic in

my attitude of trust, even if no epistemic certainty is at play, neither in the classic nor in the experiential sense.

#### 4.2 Grammar, Transmitted Knowledge, and Language

Moyal-Sharrock has argued that the security with which we hold to constitutive rules of grammar, including hinges which would actually pertain to grammar, is some kind of primitive trust or animal certainty:

Logic is seen as belonging to the realm of instinct, not reason (...), and this is reinforced by allusions to certainty as a kind of primitive (or primal) *trust*. Without this unflinching trust, there is no making sense (...). Trust, here, is not a possibility, but a logical necessity.<sup>10</sup>

The observation that our reliance upon the rules of grammar is not a conscious act but mostly done without much thinking is surely right. When I am speaking I am not thinking or choosing to trust the words I use and their meanings. I simply do. When I rely on some learned physics to do an experiment I am also (in the normal case) not choosing to trust, but trusting. However, as I was saying a moment ago this kind of non-questioned reliance, is not of the same sort as the sureness I have of “me being sitting in a chair right now,” it is no knowledge\*. The fact that I should not have it as an object of consciousness the whole time, doesn’t mean that my reliance is not based on my awareness of having sat on a chair. In the case of trusting informational states with cognitive contents, there is a difference between a) my sureness in talking hold of them and acting upon them ‘as if what they state were the case’ as if the cognitive content of such an informational state “*p*” were justified true knowledge and therefore *p* where the case, and b) my sureness of them, that is my certainty that *p*. While the instinctive attitude of trust, let’s call it ‘sureness-1,’ is the same in both cases (the personal experiential and mnemonic and the transmitted informational state cases) my certainty, or ‘sureness-2,’ in the personal case is based on knowledge\* of precisely that what I am talking hold of, but in the informational case it is not. I would reserve the notion of certainty to cases of sureness-2. If I were to be clarified in the second case and told that I am wrong, it would strike me as incomprehensible. In the first I would have much less difficulty in adapting. I could live with the sudden discovery that those that gave themselves as my parents weren’t my parents after all or that some sorts of daisies (I was so securely telling my daughter to pick up) are poisoning when in contact with salt, or that

---

<sup>10</sup> Danièle, Moyal-Sharrock, “Logic in Action: Wittgenstein's Logical Pragmatism and the Impotence of Skepticism” in *Philosophical Investigations* 26, 2 (2003): 125-148, 5

they aren't considered daisies any more, or that my security in assuming the Dutch homosexual couple in my hotel were not legally married was wrong, etc. Sureness-1, is not necessarily Sureness-2 no matter how quick and thoughtless I rely on it.<sup>11</sup> I think these two notions of 'sureness' are often conflated. I cannot be wrong in my awareness of what is standardly grouped up as a 'daisy' (as far as I know) that is, in my experience of a specimen but I could be wrong in what I take for granted daisies are like. I didn't think about the possibility of error when relying upon all that information but it didn't amount for me to knowledge\*.

One could propose that cases of transmitted knowledge should not be included among the basic hinges. Scharrock, for example, claims explicitly to be referring to what would be a universal grammar, but it is not clear to me where to draw the line between what belongs to it and what does not: not when talking about personal hinges, since she includes hinges of the kind of "those who gave themselves as my parents are such" and not when talking about transmitted informational states, which she doesn't explicitly exclude. Actually some transmitted informational states can be considered universally shared too. It seems to me that the distinction she wants to draw should be based in how relevant it is for our life. But the truly important difference relies in the fact that the ones are epistemically grounded as forms of direct knowledge\* (or directly mnemonically based upon it) while the others aren't. Furthermore, if we consider grammar as a whole, much of it, can be considered of the informational kind even if ingrained in our concepts. Is she ready to see all of this in terms of basic universal certainty of the same kind as personal experiences?

The distinction between those contingent statements that play a normative role in our frameworks and those a priori propositional truths that constitute what Wittgenstein calls 'propositions of logic' is smooth. Since the necessity of those 'truths of grammar' is due to our words' meaning, whose origin is often just as empirically contingent and taken for granted as hinges themselves are. If many of our inherited background assumptions, as Wittgenstein's own remarks show (see below), are inbuilt in language, in their previous stadium they are equally taken for granted contingent empirical statements just as hinges. Surely, functioning as word meanings they fix what are going to be the necessary rules of language, but

---

<sup>11</sup> My impression is that sureness-1 comprises both the categorical and the doxastic attitude distinctions that Scharrock (Moyal-Scharrock "Unraveling Uncertainty," 79), makes of objective certainty. Since the first of them refers, as she says, to the foundational status we give such certainties as our non-questioned inherited background, our taken for granted world picture, which comprises all that we take our world to be like in our acting, which makes no distinction between those for which we have knowledge\* and those for which we do not; it expresses also the notion of blind trust that she wants to capture through the doxastic attitude distinction.

just because we (most of us) have adopted them in the first place without much further epistemic query. That is, we might know a priori as a matter of logic that if something is 'water' it is of necessity H<sub>2</sub>O, but we are absolutely taking for granted this information, which we assume without further questioning from our chemical textbooks. That is what it means to say that they are merely a priori truths of grammar.

168. But now, what part is played by the presupposition that a substance A always reacts to a substance B in the same way, given the same circumstances? Or is that part of the definition of a substance?

Wittgenstein is quite aware of this quinean interchange in our language games, in our theories, and how this produces changes in the meanings of our words.

65. When language-games change, then there is a change in concepts, and with the concepts the meanings of words change.
410. Our knowledge forms an enormous system. And only within this system has a particular bit the value we give it.

In these passages he is considering to which extent statements like that in §168 might be conceptual truths. But many other passages in *On Certainty*, devoted to explaining hinges, include appeal to similar claims from our text books we would not question but on whose basis we go on trying to understand our world further. This is not to say that there aren't any differences between these 'logico-linguistic' statements and at least some hinges (when considering the genealogical origin of the first) since this would not apply to cases such as "my name is Ludwig Wittgenstein" or "I have never been to the moon," statements whose certainty, as is, would constitute some kind of knowledge\*. Many cases of grammar (surely not all) are, though, of the transmitted 'knowledge' sort.

In all these cases, which include much of what makes up grammar, we have no epistemic certainty but trust on rules and information that do surely have a huge epistemic import. But this trust is not detached from epistemic justification either. Not because we do as if the content transmitted is true, not because we should have the needed justification for it most of the time, but because, as Wittgenstein says, the trust itself is justified.

#### 4.3 'Grammar:' A New Foundation or a Founded One?

That much of our interaction with the world departs from such taken for granted linguistic and non-linguistic rules, no matter how *surely-1* they are relied upon, doesn't seem sufficient to accredit the trusted with a foundational role in any

meaningful sense. Being a trusted departure point does not amount to being a foundational one. What kind of a foundation would one be that instead of serving as bottom ground for revisions might be revisable itself? In the interplay between holding fast and changing the wooden planks of our ship all we might achieve is indeed just a form of coherentism. In reading Scharrock's<sup>12</sup> paragraphs on the 'historically or humanly constrained a priori,' which she describes as providing some non-traditional 'logical necessity with a human face,' one gets the impression, though, as if she would be presenting this linguistic grammar-bounded necessity as some sort of last foundation beyond which we could not go. This is not to deny that grammar might change, she agrees, but this would be "all the necessity we can get," a necessity "without absolutes." In providing this picture she appeals to Cavell's reading of Wittgenstein and offers an interpretation reminiscent of McDowell's *Mind and World*. Suggesting that beyond our linguistic apprehension all there might be are *causal* constraints of which we can have no cognitive understanding but just neural impingements, no perceptual awareness when not conceptual. This clearly reminds of McDowell's critique of Davidson. Even if there should be some causal conditioning at a neural level it has no cognitive import and therefore our perception must be understood as conceptual all the way down.

It is true that when Scharrock talks about foundations she is most of the time referring to 'universal grammar' and many of those hinges she appeals to are of what I call 'the experiential kind;' although, as I said, there is no clear separation between these and those of the transmitted knowledge kind. Regarding the experiential ones she considers, it is unclear to me how she puts together the kind of 'animal certainty' she attributes to our ineffable reliance upon such hinges, with her idea that the last ground is some kind of conceptually apprehended 'historical a priori.' Since the certainty with which animals rely on there being a world with a given structure while moving around in it can just be of the experiential non-conceptual kind or due to causal conditioning. While ours, even if understood as ineffable, must be of the propositional conceptual sort if we are to make it cohere with the picture suggested in the paragraphs on the 'historical a priori.' But maybe the comparison was just metaphorical with no pretension of similarity beyond the immediacy aspect of our reactions.

On the other hand, when talking about the linguistic a priori statements of grammar it won't be easy either to sort out those linguistic statements that are to be seen as universal from those that are not. At least I don't think this is

---

<sup>12</sup> Moyal-Scharrock "Logic in Action," 12-13

achievable without going precisely beyond the “historical a priori” and into the absolute position she attempts to avoid.

As opposed to this I do think that we can do better than holding securely to some contents but denying epistemic certainty and any deeper foundation to those (however complex) linguistic a priori statements of grammar. Our security relies not in the immediacy of our trust, nor the necessity to hold upon the transmitted, but on there being, as Wittgenstein says, a justification for the trust itself; one that at the very bottom grounds the same way that our personal experiences do. Because, even if, as Scharrock quotes Wittgenstein in saying, “language is not based on a ratiocination or epistemic agreement” – not in the conscious sense of a ceremony of learning or the signing of a covenant, not any more than the ‘Social Contract’ for accepting common norms can be said to be – it is not disconnected of its epistemic background of origination either. Actually conceptual trust, when conscious of its origins, is trust in there being along with our historical tradition some kind of epistemic background to our conceptual discriminations on whose basis we can rely when it should not be obvious of itself.

168. (...) But in the end I rely on these experiences, or on the reports of them, I feel no scruples about ordering my own activities in accordance with them.  
– But hasn't this trust also proved itself? So far as I can judge – yes.

Sometimes it might be more a matter of conceptually registered or institutionalized practices having pragmatic sense but also here we directly see, or presume worthy, the purpose for which they are established. That is, we surely do not ‘know’ but take much for granted and trust while assuming that mostly there is some kind of epistemic justification (be it: research done, discovered connections, properties, attributes or vulnerabilities, or mere relevant discriminations, relations, affections, etc. whose distinction justifies our use). But we knew too how to rewind back and prove the genealogical epistemic lineage of our words, their purpose and legitimacy and, if necessary, how to go about to change them. Actually anything that might count as the evolution of human knowledge is a matter of doing precisely that. Even in the more pragmatic cases we can come to see that new or conflicting purposes require expansion or restriction of corresponding word meanings (be it ‘marriage’ or ‘phablet-phones’). Therefore, a better picture of our trust might be something like a ‘passing the torch’ of the epistemic trustworthiness of our forbearers and conditional to the rightness of their achievements.

The only kind of foundation I can think of in the light of these reflections is therefore one of the traditional sort. That is, one that reassesses the role of a world

we can perceive, we can 'be familiar with'<sup>13</sup> (in a cognitive and more than causal sense) previous to our conceptual apprehensions; a world that possess the needed sort of autonomy and universality to correct our meanings, frustrate our purposes and resolve disputes.

### 5. Moore's 'Recognitional Statements:' Responding to the Sceptic

Wittgenstein reflections approach the problem from a different perspective in considering why Moore's claim that "he knows that this is a hand" adds nothing to his claim that "this is a hand" and is no help against the sceptic. Wittgenstein appeals here to the use of words in language games. "This is a hand" or "a tree" because that is the way our words are used, that is what we call 'a hand.' Here the point is not one about grammar but about what our words are applied to. Now, these claims could be made no matter the type of words we use if we use them properly, whether we talk about 'hands,' 'neutrinos,' 'generous' people or 'unchaste women.' In a sense in all these cases we have the certainty (at least in standard cases) that that is the way words are used in our language games. The "I know" adds of course no further assurance that would not be there in what makes us issue the statement in the first place. So, quoting Williamson again, the knowledge (or *knowledge*\*) is in that first step already, in the claim that *p is the case*. That perfectly coheres with what Wittgenstein says. But if we push this point further making it say that the 'doing it that way (rightly)' gives us an assurance of the truth of what *our words say*, we might seem to land here again in some kind of direct realism of the mcdowellian sort. There are, however, substantial remarks that show that this is not how Wittgenstein means it. Consider this:

584. Would it be possible to make use of the verb 'know' only in the question "How do you know?" following a simple assertion? – Instead of "I already know that" one says "I am familiar with that;" and this follows only upon being told the fact. But what does one say instead of "I know what that is?"

Here the difference between being 'familiar with' something and knowing 'what it is,' are two pair of shoes. Even if that is how we use our words, the trust is not necessarily an identification between the correct use of words and the acquainted experience that makes it correct. What I trust is that this is the way to prove whether this is 'a hand' or that a 'neutrino' in my language game. But again our notion of 'neutrino' might perfectly evolve and as a consequence our claims

---

<sup>13</sup> See Wittgenstein's *On Certainty*, remark 584 quoted further down in this paper



about what the sorted out items in its extension are (or signalize in this case) too<sup>14</sup>. The point is not one about being certain of the epistemic adequacy of our conceptual terms and corresponding truths and thereupon presumed facts, but rather a certainty in the way we prove that our conceptual conceptions apply, and in my experience of having done so (in terms of knowledge\*); which can be equally expanded to a certainty in our way to determine what counts as correcting them if required. Both the knowledge that this is how my concepts are applied and that I am before an application case are forms of experiential or experientially derived knowledge\*. I am certain (sureness-2) that this is the way to prove whether a cell is an 'eukaryote,' I am certain (sureness-2) that I have the experience necessary to prove it, but I simply trust (sureness-1) that the distinctions made with the term in my biological text books are exhaustive enough. I could perfectly accommodate again that against what I was told there might be eukaryotes without mitochondria organelles, as there are<sup>15</sup> and my textbooks have to be changed. Again, I accept the change because I trust that scientists made the necessary experiences to prove it.

It is from this perspective, of course, that the question of certainty in language use and in experiential hinges become connected. In both cases it is the experiential basis that determines certainty (sureness-2).

Actually, my impression is that despite the heterogeneous appearances of Wittgenstein's differing approaches and the varied layers of depth he considers, they all trace together the intended territory that demarcates what he was after. A better guidance to a unified interpretation of what this is requires asking ourselves in which sense all these varied remarks could deliver an answer to the sceptic. This approach can give us a sense of the way in which we might speak of a foundational ground to them all.

---

<sup>14</sup> This is no concession to a metaphysical realist or referentialist position. The referential set is always classified according to some common descriptive properties, and if these descriptions (as a result of experiencing with given, possible and frontier set members) require modification, expansion restriction or whatever, the referential class can be modified too or even be dissolved completely as such. The relation between referential class and description operates both ways.

<sup>15</sup> Recent biological research has found that *Monocercomonoides* Sp lacks mitochondrial organelle, what was considered essential for eukaryotes, but are in all other relevant respects an eukaryote. See [http://www.cell.com/current-biology/fulltext/S0960-9822\(16\)30263-9](http://www.cell.com/current-biology/fulltext/S0960-9822(16)30263-9). This is an information so basic that it might require a reformulation of one of the most assented claims in our biological textbooks.

## 6. A Unified Account

The reason why we might come to the conclusion that hinges do not fit the bill of a foundational account is, I think, because we fix our attention in a given set of statements such as “My name is L.W,” “I have never been to Asia Minor,” “This is a hand,” etc. that are quite heterogeneous. In addition the phenomenon of certainty or trust that is each time considered is not homogeneous either. We find disparage claims: we are told that some such certitudes cannot be questioned because doing so would destroy our world picture, we would be considered insane and so on; other remarks advising us that the point is not that we should be infallible about many such claims but that it would be nonsensical that they all were, some pointing at the wider use of linguistic and non-linguistic statements as constitutive rules of language games, some appealing to statements of our inherited gnoseological patrimony, to contextual and culturally relative ones and so on.

If we look at things, though, from the standpoint of giving an answer to the sceptic, it becomes increasingly clear that it is not a given set or a given class of statements that are beyond doubt or could be considered foundational. It is rather the fact that we can trust for example, not primarily always the adequacy of our specific conceptual tools or a priori rules (in discourse or action) but rather, as seen, that they are proven a given way. That is, that the way we ascertain ourselves of their correctness, on the basis of both our perception and reliance on the existence of meanings as a whole, is non-negotiable, something without which our world does break down. In the same way, it is not specifically that someone could not in some instance be wrong about his name, or must hold it fast, or that I have never been to China or the Moon or that there are some stairs outside my bedroom etc. It is rather the fact that we must rely on our experiences and our memory that without the reliance on this capacity to store and recall our experiences (especially those that draw the central ‘files’ or central categories of our self-conception: who I am, where have I been, what have I done, what I am up to) of course nothing would be as it is, no world at all would be given to us. This is exactly what the sceptic is questioning, the very basis of our knowledge, the conditions of possibility of our world being what we take it to be, both the cognitive conditions of the knowing subject and the conditions of the known world; the sceptic is questioning whether our trusted apparatus of cognition, our memory, our use of words and the evidential basis that support and confirms our claims could not be wrong. It is the very fact that we rely on such means for knowledge that is disputed. Other seemingly dissimilar claims like: “no one has ever been to the moon” but also the remarks on ‘chemistry or anatomy’ allow an

interpretation on the same lines without difficulties. “No one has ever been to the moon” also touches upon some basic issue about our world constitution and our knowledge structures, since the reason why this is not believable is because in order for this to be true at all, the means that would make it possible must have been achieved and they aren’t. Since for us there are means for anything that could count as knowledge, not anything counts as confirmation or disconfirmation. Just as there are ways to prove that we trust, there are ways that are beyond our possibilities as human beings. The same way we do trust, we rely upon the information that our textbooks, or ancestors, have transmitted us, but we do because, as Wittgenstein says, this is justified. We think it was justified in its origins through the same procedures all our own knowledge\* experiences are. We think someone was in an epistemic position to acquire the needed experience and acquire the claimed knowledge with human means or correspondingly enhanced technical possibilities.<sup>16</sup> Because, at the very end, transmitted knowledge too, even if we should not have experiential certainty (sureness-2) about it, must have been acquired on the basis of some such experiential certainty by our ancestors. Those are the conditions of human knowledge and, therefore, those of our trust too.

108. “But is there then no objective truth? Isn’t it true, or false, that someone has been on the moon?” If we are thinking within our system, then it is certain that no one has ever been on the moon. Not merely is nothing of the sort ever seriously reported to us by reasonable people, but our whole system of physics forbids us to believe it. For this demands answers to the questions “How did he overcome the force of gravity?” “How could he live without an atmosphere?” and a thousand others which could not be answered. But suppose that instead of all these answers we met the reply: “We don’t know how one gets to the moon, but those who get there know at once that they are there; and even you can’t explain everything.” We should feel ourselves intellectually very distant from someone who said this.

670. We might speak of fundamental principles of human enquiry.

Similarly see  $\zeta 671$ <sup>17</sup> on this point. Without rackets and the necessary astronomic and physical knowledge any such experience was impossible. The

---

<sup>16</sup> This reminds on how William Kingdon Clifford, “The Ethics of Belief” in *Contemporary Review* (1877), at present in. *The Ethics of Belief and other essays* (Amherst: Prometheus Books, 1999), explains when trust in authorities and transmitted knowledge is justified: ultimately, he says, when we know that someone could have been in the required position, the information could have been acquired with human means etc. Of course, there are also pragmatic reasons to go on trusting when the trusted seems to fit well with our world.

<sup>17</sup> Wittgenstein’s *On Certainty*,  $\zeta 671$ , I fly from here to a part of the world where the people have only indefinite information, or none at all, about the possibility of flying. I tell them I have

remarks on chemistry or anatomy appeal again to the role played by our experience of acquisition and our memory in the transmission of knowledge and ultimately to the link of our trust to its foundation. The role of memory too in the building of a historical narrative, both personal and socio-cultural, is essential to the temporally structured human beings that we are: to our self-understanding and our moving around in a trustable environment, to the conception of the world we take ourselves to live in. There is an homogeneity in what all these aspects point to. It is, thus, not to be searched for in the specific remarks but in the capacities and ways through which we support those claims.

In this line the real response to the sceptic is then that coming to realize the very possibility he envisions, the possibility of being wrong about the way we prove and ascertain our conception of the world requires precisely making use of those very capacities he is questioning with no more presumption of certainty than what we already have. There is no going beyond it for a human being, sceptic included. Maybe the answer is not that it is impossible but that it is impossible for us to ever know it, not because we could not awake from a common dream or find much of our experience illusory, but because doing so delivers us back in exactly the same situation we were before with exactly the same kind of warranty and ways of proving.

301. Supposing it wasn't true that the earth had already existed long before I was born - how should we imagine the mistake being discovered?

641. "He told me about it today - I can't be making a mistake about that." - But what if it does turn out to be wrong?! - Mustn't one make a distinction between the ways in which something 'turns out wrong'? - How *can it be shown* that my statement was wrong? Here evidence is facing evidence, and it must be *decided* which is to give way.

Illustrative to this point also paragraphs ζ642, ζ650.<sup>18</sup> So the things we are certain of are those regarding the very conditions of possibility of our knowing (or

---

just flown there from... They ask me if I might be mistaken. - They have obviously a false impression of how the thing happens. (If I were packed up in a box it would be possible for me to be mistaken about the way I had travelled.) If I simply tell them that I can't be mistaken, that won't perhaps convince them; but it will, if I describe the actual procedure to them. Then they will certainly not bring the possibility of a mistake into the question. But for all that - even if they trust me - they might believe I had been dreaming or that magic had made me imagine it.

<sup>18</sup> ζ642. But suppose someone produced the scruple: what if I suddenly as it were woke up and said "Just think, I've been imagining I was called L.W.!" - well, who says that I don't wake up once again and call *this* an extraordinary fancy, and so on?

ζ650. This surely means: the possibility of a *mistake* can be eliminated in certain (numerous) cases. - And one does eliminate mistakes in calculation in this way. For when a calculation has

correcting) something, which relates to a conception of the world and ourselves. Now, can we pronounce what these conditions are, so to speak, can we state them?

I think that some other remarks of Wittgenstein, some of his basic hinges (those shown under a) at the beginning,<sup>19</sup> do give voice to this kind of fundamental assumptions about the world:

- 101. Such a proposition might be e.g. "My body has never disappeared and reappeared again after an interval."
- 134. After putting a book in a drawer, I assume it is there, unless... "Experience always proves me right. There is no well attested case of a book's (simply) disappearing."...
- 153. No one ever taught me that my hands don't disappear when I am not paying attention to them. Nor can I be said to presuppose the truth of this proposition in my assertions etc., (as if they rested on it) while it only gets sense from the rest of our procedure of asserting.
- 234. I believe that I have forebears, and that every human being has them.
- 234. ...I believe that the earth is a body on whose surface we move and that it no more suddenly disappears or the like than any other solid body: this table, this house, this tree, etc. If I wanted to doubt the existence of the earth long before my birth, I should have to doubt all sorts of things that stand fast for me.

Most of these remarks are about the permanency of objects, of solid bodies in space throughout time. Again, these are not any particular empirical statements that I hold fast but the very notion of what our conception of a solid body is.

Two points here: first, as just stated, it does belong to our notion of a solid body that it extends in space and remains in time whether we perceive it or not. It is an empirical statement, yes, but as I was suggesting before, because our concepts are at the very end grounded (at least to some extent) on empirical apprehensions. Second, we could say that what is being 'put into words' here (while normally just taken for granted) has more to do with our expectations about the existent world than with what we take the world to be from a more phenomenological perspective. With the constitution of entities as such (those we 'can be familiar with' or not) more than with what the entities are in terms of 'a tree,' 'a house' etc. From a more classical perspective we would say that it somehow relies in what

---

been checked over and over again one cannot then say "Its rightness is still only *very probable* – for an error may always still have slipped in." For suppose it did seem for once as if an error had been discovered – why shouldn't we suspect an error *here*?

<sup>19</sup> See *supra*, p. 73.

classical authors tried to capture through the notion of ‘categories.’ Those aspects that pertain to any physical body whatsoever, its being a differentiated and numerical (quantitative) entity, it being substantially extensive, it remaining in time and in a given location when not seen etc. it's possessing some quality or other. Many of the commentaries that Wittgenstein makes refer at the very end to these aspects, which require eventually not only experiences but results out of inferential processes: If I can retain a memory of myself in space before this moment, there must have existed such a space before. Knowing that you are in a given location (and therefore not in China),<sup>20</sup> that objects don't disappear, that if something is in a given location and not moved it remains there, the relations of an entity to others around it in its spatial setting (the stairs outside my room) or temporal coordinates:<sup>21</sup> that you are living in a given moment in time and there was time before and after, that you did something this morning. As Ackrill<sup>22</sup> says about the problem Aristotle tries to capture through his *Categories* “it is not primarily or explicitly about names, but about the things that names signify...” It is like an attempt to put into words what we need to take for granted for the world we talk about to be the world we refer to and confirms our statements and, also, what we ourselves have to be like to register this world the way we do, our registering succession of events through memory, our using words with meanings to describe it.

This spatio-temporal character of human beings, both as entities ourselves in a given location related to other entities and as cognitive beings, having a representational structuring capability (that from an epistemic perspective reproduces this order in understanding and situates itself in it) can be considered as pertaining to our notion of ‘a human (rational) being’ and thus necessarily true of it. I think these are the kind of phenomena that Wittgenstein, in line with Aristotle and with Kant, was in his own peculiar way after.<sup>23</sup>

Maybe what raises much puzzlement is that Wittgenstein brings this problem further than his predecessors. Since he would seem to include in it a more historical aspect too.

---

<sup>20</sup> At least in the macrocosmic world of experiences, that quanta should have the capacity to be in two places at the same time I find difficult to digest but won't, of course, dare to dispute. I align myself with both trust and puzzlement.

<sup>21</sup> Some commentaries in Wright, “Skepticism, Certainty,” would seem to go along these lines.

<sup>22</sup> John Ackrill, *Aristotle, Categories and De Interpretatione* (Oxford: Clarendon Press, 1963), 71

<sup>23</sup> Actually, the response to the sceptic considered before would cohere somewhat with a Kantian epistemic approach since any discovery of a mistake must be for Kant too stated as a mistake in terms of a phenomenal world beyond which there are nor mistakes nor phenomena but inconceivable noumeno.

99. And the bank of that river consists partly of hard rock, subject to no alteration or only to an imperceptible one, partly of sand, which now in one place or in another gets washed away, or deposited.

His going into the structure of language games, where some knowledge is taken for granted while we go on playing it and then again it gets ‘washed away,’ our trusting knowledge and (changing) background assumptions appears to put the problem on a different level. But even if what we do in each single case might change, the background assumptions being removed, the fact that we do and must trust in such a way, that we must depart from acquired knowledge by our forebears (mnemonically transmitted, justifiably trusted) and take history into account is a universal one about ourselves too, about our existential historically bounded self. As such, though, it can be considered a stable structure of our being too.

It has sometimes being claimed that the phenomenological rediscovery of our finite historical character, our being subsumed in the current of history and its changing moves, already there in Hegel, somehow superseded previous transcendental approaches that focused more on what are the cognitive capabilities of the knowing subject and the world to be known. That once the inter-subjective socio-historical and pragmatic character of language was discovered, the idea of a transcendental conceptualization, a categorical (absolute?) structuring of the world (and ourselves) had to be abandoned and there was no going backwards. However, what we might find here is a realization that both can and must be put together in an integrative picture, instead of the one being replaced by the other, to show what we actually and universally are. To answer the sceptic we encountered what is it that we must take for granted if our knowledge is to be possible at all: the epistemically (cognitive though not linguistic) accessed evidence that corroborates (and corrects) our statements (and concepts), upon which the sceptic himself must rely to show us wrong (beyond which a noumeno relies), our own capacities to register, locate, relate, remember, and talk about through developed meanings and inferential connections; and, also, transcendental historically bounded structures: our having to rely in a mnemonically transmitted narrative of historical knowledge and language (not of course its possibly changing content each time) about the way the world is and we are in order to have a world and a self-understanding; our having to rely on it in our living, even if it changes because at the very end all waters come from the same source (even if our hypothesizing and elaborating upon them might not) that all is intrinsic to our natures as well. As Wittgenstein would put it: There is something universal here too!





## **DISCUSSION NOTES/DEBATE**



# METHODS MATTER: BEATING THE BACKWARD CLOCK

Murray CLARKE, Fred ADAMS, and John A. BARKER

ABSTRACT: In “Beat the (Backward) Clock,” we argued that John Williams and Neil Sinhababu’s Backward Clock Case fails to be a counterexample to Robert Nozick’s or Fred Dretske’s Theories of Knowledge. Williams’ reply to our paper, “There’s Nothing to Beat a Backward Clock: A Rejoinder to Adams, Barker and Clarke,” is a further attempt to defend their counterexample against a range of objections. In this paper, we argue that, despite the number and length of footnotes, Williams is still wrong.

KEYWORDS: backward clock, Fred Dretske, knowledge, Robert Nozick, Neil Sinhababu, tracking theories, John Williams

In “Beat the (Backward) Clock,”<sup>1</sup> we argued that John Williams and Neil Sinhababu’s Backward Clock Case fails to be a counterexample to Robert Nozick’s or Fred Dretske’s Theories of Knowledge.<sup>2</sup> Williams’ reply to our paper, “There’s Nothing to Beat a Backward Clock: A Rejoinder to Adams, Barker and Clarke,”<sup>3</sup> is a further attempt to defend their counterexample against a range of objections. In this paper, we argue that, despite the number and length of footnotes, Williams is still wrong. As Shakespeare might have opined: “The Man doth protest too much, methinks!” Tracking theories still beat the clock!

The central issue at the heart of our disagreement with Williams and Sinhababu is the role that methods (Nozick) or reasons (Dretske) play in these accounts of knowledge. Tracking the truth crucially depends on the method or reasons employed in the acquisition of belief. So central are such methods or reasons that it makes little sense to talk about beliefs being sensitive or adherent to the truth except in the context of the method or reasons employed to arrive at such beliefs. It is Williams and Sinhababu’s failure to accord method or reasons

---

<sup>1</sup> Fred Adams, John A. Barker, and Murray Clarke, “Beat the (Backward) Clock,” *Logos & Episteme* VII, 3 (2016): 353-361.

<sup>2</sup> See John N. Williams and Neil Sinhababu, “The Backward Clock, Truth-Tracking, and Safety,” *Journal of Philosophy* 112, 1 (2015): 46-55. In this article they cite Adams and Clarke’s earlier defence of tracking theories in “Resurrecting the Tracking Theories,” *Australasian Journal of Philosophy* LXXXIII, 2, (2005): 207-221.

<sup>3</sup> John N. Williams, “There’s Nothing to Beat a Backward Clock: A Rejoinder to Adams, Barker and Clarke,” *Logos & Episteme* VII, 3 (2016): 363-378.

their proper place in the respective accounts of knowledge that creates the illusion that the Backward Clock Case is a counterexample. It is no such thing. In what follows, we first revisit Nozick's account of knowledge and the role that methods play in his account. Later, we show why Williams' latest response fails to faithfully respect the tenets of Nozick's view. Finally, we discuss Dretske's appeal to reasons and explain how this appeal does the work that methods do for Nozick. The result is that Williams' response also fails to address Dretske's actual theory.

### 1. Nozick's Analysis of Knowledge

In fact, there are three accounts of knowledge that Nozick provides: a simplified tracking account where he claims method need not be mentioned because it is not relevant for some straightforward cases, a methods account to deal with more complicated cases where single methods are at issue, and an outweighing account that involves two or more methods of arriving at belief. For our purposes, it is the second account that is needed to respond to Williams' response to us. It states that:

S knows, via method (or way of believing) M that  $p$  iff:

1.  $p$  is true.
2. S believes, via method or way of coming to believe M, that  $p$ .
3. If  $p$  weren't true and S were to use M to arrive at a belief whether (or not)  $p$ , then S wouldn't believe, via M, that  $p$ .
4. If  $p$  were true and S were to use M to arrive at a belief whether (or not)  $p$ , then S would believe, via M, that  $p$ .<sup>4</sup>

Notice that the truth-tracking sensitivity and adherence conditions, i.e., 3 and 4, both explicitly refer to the method. Here is what Nozick says about method:

We need to relate this technical locution to our ordinary notion of knowledge. If only one method M is actually or subjunctively relevant to S's belief that  $p$ , then, simply S knows that  $p$  (according to our ordinary notion) if and only if that method M is such that S knows that  $p$  via M. Some situations involve multiple methods, however...<sup>5</sup>

Nozick goes on to discuss the Father/Son Court Case in order to introduce the third, outweighing, account of knowledge where multiple methods are at play. We now turn to Williams' putative counterexample to Nozick's account of knowledge.

---

<sup>4</sup> Robert Nozick, *Philosophical Explanations* (Cambridge: Harvard University Press, 1981), 176.

<sup>5</sup> Nozick, *Philosophical Explanations*, 176.

## 2. Williams' Response to Our Response to Backward Clock

Williams and Sinhababu describe The Backward Clock Case as follows:

You habitually nap between 4 p.m. and 5 p.m. Your method of ascertaining the time you wake is to look at your clock, one you know has always worked perfectly reliably. Unbeknownst to you, your clock is a special model designed by a cult that regards the hour starting from 4 p.m. today as cursed, and wants clocks not to run forwards during that hour. So your clock is designed to run perfectly reliably backwards during that hour. At 4 p.m. the hands of the clock jumped to 5 p.m., and it has been running reliably backwards since then. This clock is analogue so its hands sweep its face continuously, but it has no second hand so you cannot tell that it is running backward from a quick glance. Awaking, you look at the clock at exactly 4:30 p.m. and observe that its hands point to 4:30 p.m. Accordingly, you form the belief that it is 4:30 p.m.<sup>6</sup>

They argue that all of Nozick's conditions are satisfied concerning this example but that it fails to be a case of knowledge, thus demonstrating that Nozick's conditions are too weak for knowledge. In particular, if the time were not 4:30 p.m. (not *p*) then you would not believe that it was because, for instance, at 4:31 p.m. you would believe that it is 4:29 p.m. and so forth. You would hold a false belief but you would succeed in not believing that it was 4:30 p.m. as required by Nozick's account. Similarly, in other circumstances where it were 4:30, (say, you were closer to the clock) you would believe that it is 4:30 p.m. Hence, the belief is both sensitive to, and adherent to, the truth value of *p*, i.e., 4:30 p.m., and so satisfies Nozick's truth-tracking conditions, 3 and 4. Now Williams is correct to contend that we have challenged the claim by Williams and Sinhababu that Nozick's third condition is satisfied in the Backward Clock Case. We certainly do deny that Nozick's third condition is satisfied here. Why? This is because we think that the method, i.e., 'looking at the clock and determining what it says,'<sup>7</sup> is too equivocal to yield knowledge. As we pointed out, on the most plausible interpretation of the example the clock was designed by the cult clockmakers to fool people during the cursed hour. In effect, the clock 'lies' by displaying, for instance, 4:35 when the time is 4:25 and vice-versa. But 'Ted,' as we called the clockmaker, wasn't a perfect liar, for when the clock displays 4:30 it is saying that the time is 4:30. Given Ted's deceitful intentions, the clock might have said the time was 4:30 even if it hadn't been 4:30. For instance, as we suggested, Ted might have made the clock run slowly all during the cursed hour. Hence, condition 3 is not satisfied.

---

<sup>6</sup> Williams and Sinhababu, "The Backward Clock," p. 48.

<sup>7</sup> Adams et.al., "Beat the (Backward)," 355.

At all other times during the hour between four and five, however, the clock succeeds in lying successfully. Hence, for any time during the hour, the method of 'reading what the clock displays' will generate false beliefs. As such, the method employed will generate false beliefs and so your beliefs will be insensitive to the truth-value of  $p$  for all values of  $p$  other than 4:30. But we also think that Nozick's fourth condition, the adherence condition, is not satisfied in the Backward Clock Case. It is true that we accept that the cognizer has a true belief at 4:30 p.m. But the cognizer does not satisfy the adherence condition. Why? This is because satisfying the adherence condition requires something much stronger than mere true belief, it requires that if it were 4:30 (in other circumstances) then one would believe that it was 4:30 p.m. But this is exactly what is not the case with the Backward Clock. As we pointed out: "His belief is that it is 4:30, and it happens to be 4:30. But it is not the case that he believes it is 4:30 because it is 4:30-his believing it to be 4:30 is not explained by the fact that it is 4:30."<sup>8</sup> The signal is too equivocal to be reliable in other circumstances since the clock might have been made not to read 4:30 even if it was 4:30 and so one would not believe that  $P$  though  $P$  is true. Suppose, for instance, the clock shuts off at 4:30 for one minute but otherwise reliably runs backwards from 5 until 4. You wake up, look up, see no time on display, and suspend judgement on the time for that minute. In such a possible world,  $P$  is true but you don't believe that  $P$  and so condition 4 is not satisfied. We deny, therefore, that condition four is satisfied concerning the Backward Clock Case.

That the method generates a true belief at 4:30 is a chronometric accident caused by the mistake that Ted made in the construction of the clock. Now it is here that Williams suggests that the method need not be reliable for all of these other times during the hour as long as the method works for 4:30 then conditions 1-4 will be satisfied and Nozick's account will incorrectly generate the result that one knows that  $p$ . Our mistake, on William's view, is mistaking Nozick's sensitivity condition as a constraint on **METHOD** rather than **BELIEF**. But it is Williams who misunderstands Nozick's theory, not us. This is because it is exactly the **method** that ensures that the correct connection between belief and fact obtains when we know some factual belief, that the belief is both sensitive and adherent. That method must be absolutely reliable with respect to a variety of input beliefs in near possible worlds for  $S$  to know that  $p$ . It cannot be reliable for just one belief in near possible worlds, such an equivocal method would fail to deliver the epistemic goods. Consider Nozick's Grandmother Case: she believes that her Grandson is well by appeal to visual perception, but if he were sick then

---

<sup>8</sup> Adams et.al., "Beat the (Backward)," 359.

her daughter would tell her that he is well anyway. But Nozick tells us that that alternative method, i.e., testimony, should not be allowed to undermine the knowledge caused by the method of visual perception that the Grandmother has. It is exactly because the Grandmother's method, i.e., visual perception, is absolutely reliable for close distances to her in near possible worlds, that she can confidently say that her Grandson is well. Nozick was not just saying that this belief was true in the actual world but subjunctively true as well. That subjunctive truth implicitly refers to what the Grandmother would have believed about  $p$ , for all  $p$ , that are cases of visual perception under such circumstances for that Grandmother. What provides the basis for the tracking of truth in near possible worlds for that belief, i.e.,  $p$ , is exactly the fact that that method would track the truth for any visual perceptual belief for the Grandmother under those circumstances in near possible worlds. To think otherwise, is to deny the fact that we are talking about tracking accounts of knowledge at all!

Here is what Nozick says about the Grandmother Case:

Recall the grandmother who sees her grandson visit her and so believes he is healthy and ambulatory; yet if he weren't ambulatory, other relatives would tell her he was fine to spare her anxiety and upset. She sees her grandson walking; does she know he is ambulatory? According to condition 3 we must ask what she would believe if he weren't ambulatory. If the method via which she believes is not held fixed, the answer will be wrong. True, if he weren't ambulatory, she would then believe he was (via hearing about him from other relatives). But the relevant question is: what would she believe if he weren't ambulatory and (as before) she saw him and spoke to him. Thus, to reach the correct answer about her knowledge, the method must be held fixed—that is one of the reasons why we introduced explicit reference to the method or way of believing.<sup>9</sup>

Nozick makes it explicit that the Grandmother would generate other beliefs about her Grandson that are veridical in near possible worlds. The whole mechanism of possible worlds is simply a device to talk about other beliefs of the same kind, i.e., visual perception beliefs, and insist that the method must be held fixed: all visual perception beliefs, all beliefs of that type about her Grandson, must be reliably produced by that method for the Grandmother under those circumstances in order for her to know that  $p$ . That is what is meant by saying that a belief is not only true, but subjunctively true: one would have arrived at other related visual perception beliefs veridically in near possible worlds, one would have gotten things right in near possible worlds. Hence, the idea that one must track the truth of 'not  $p$ ' veridically demonstrates that the method must be

---

<sup>9</sup> Nozick, *Philosophical Explanations*, 216.

absolutely reliable for beliefs of that type. After all, the **very idea** of ‘not  $p$ ’ includes, among other things, the entire universe of visual perception beliefs other than  $p$ ! When Nozick asserts condition three his point is that, given your method, you would track the truth of ‘not  $p$ ’ **veridically**, whether the resulting belief be  $q$ ,  $r$ , or  $s$ . Hence, the Grandmother might form a different belief about her Grandson, such as  $q$ : “My Grandson is ill as evidenced by the fact that he cannot walk.” One must not believe that  $p$  and veridically track the truth such that you don’t believe that, say,  $q$  falsely. ‘Not  $p$ ’ is a label for a universe of possible beliefs that one would reliably get correct in near possible worlds. That is what provides the subjunctive strengthening of the causal condition that Nozick felt was needed. He, like Dretske, felt that the causal condition was correct as far as it went, it just did not go far enough. The subjunctive allowed them to talk not just about getting the particular causal connection between belief and a fact right in the actual world, but getting the belief/fact connection right in near possible worlds as well. It strains credulity beyond the stratosphere to think that Nozick intended ‘not  $p$ ’ to include the idea that one might latch on to ‘not  $p$ ’ beliefs that were false and still be using a reliable method. Tracking the truth presupposes the **reliability** of the **method** for producing truth. Reading his account in any other way is simply a misreading of Nozick. The sensitivity condition, condition three, builds into it the idea of an absolutely reliable method, in near possible worlds but not all possible worlds, for tracking the truth of ‘not  $p$ .’ The reliability of the method concerning  $p$  is substantiated by the fact that that method would be reliable, in near possible worlds, for generating true beliefs. This is the correct understanding of what Nozick was arguing for. The Backward Clock Case would have led to a quick death for Nozick’s theory within minutes of his thinking of the theory if he had understood his own theory as Williams does. Why? Because Nozick himself would have understood that the theory, understood in that way, was bankrupt! But, says Williams, mightn’t Nozick not have noticed that the theory has this very odd consequence that even false ‘not  $p$ ’ beliefs can serve to confirm condition three of the theory for a particular  $p$ ? The correct answer to this objection is: No, only an extremely uncharitable reading of the theory could possibly interpret Nozick as intending, or leaving open, or suggesting, or not noticing, this interpretation of the theory. The principle of charity counsels us to avoid implausible and unlikely interpretations of the words of an author. If ever there was an implausible and unlikely reading of a theory, Williams’ reading of Nozick is it. There is no possible world where this reading of Nozick passes muster!

The ‘Boy who cried Wolf’ Case, which we developed at length in our reply, was expressly devised to make the point that equivocal signals will not generate



knowledge. Conveniently, Williams ignores this argument from our response. As we noted:

For any time other than exactly 4:30, the subject's belief during that hour-long period will be false. Why? Because the clock lies for all but one moment during that hour-long period. And worst of all, there is nothing in the signal sent by the clock to differentiate when it is telling the false time from when it is telling a true time.

This should remind one of the "little boy who cried 'wolf.'" The boy cries 'wolf' over and over when there is no wolf. Then on the one occasion when there is a wolf and he cries 'wolf,' his cry has become to equivocal, no one can tell from his cry that a wolf is actually there on that one occasion. His cry of 'wolf' still means wolf, but it does not carry the information that there is a wolf. Similarly, the clock's face emits false testimony for 59 minutes during that hour from 4:00 to 5:00.<sup>10</sup>

The appeal to Shannon's information theory in *Knowledge and the Flow of Information* was the way that Dretske chose to instantiate the subjunctive condition that he imposed from his early Conclusive Reasons Account of knowledge. The role that Dretske's subjunctive conditional played in his Conclusive Reasons Account of knowledge, i.e., Given R and fixed circumstances C, it is not physically possible that not p, in turn, was taken over by Nozick's third condition on knowledge. Nozick's subjunctive conditional is not identical to Dretske's subjunctive conditional but it imposes a similar constraint on knowledge.<sup>11</sup> At any rate, our appeal to information in this example was intended to draw out the problem in Williams' Backward Clock Case by appealing to a Dretskean notion of information via the Boy Who Cried Wolf Case. The upshot is that neither the interpreter of the Boy who cried Wolf nor the ordinary person who sees the Backward Clock displaying 4:30 is in a position to know anything about the wolf, or the time, in such cases.

Williams also mentions that: "Nozick introduces methods into his analysis, not as a way of elucidating sensitivity, but in order to avoid a counterexample."<sup>12</sup> While this is one reason to introduce methods, it is hardly Nozick's only reason to do so. As we mentioned in an earlier paper,<sup>13</sup> how could condition (4) be satisfied by anyone, if methods aren't the means? Truths don't just pop into heads. It often

---

<sup>10</sup> Adams et.al., "Beat the (Backward)," 358.

<sup>11</sup> Nozick, in speaking of Dretske's condition two, says that: "While this condition corresponds to our condition 3, he has nothing corresponding to 4." (Nozick, *Philosophical Explanations*, 689, footnote 53.)

<sup>12</sup> Williams, "There's Nothing," 364.

<sup>13</sup> Adams and Clarke, "Resurrecting," 214.

takes hard work (science, detectives) to discover them. We take it as an obvious fact about tracking theories (Nozick's or Dretske's) that beliefs only track in virtue of reasons or methods. Otherwise, such theories would make no sense. In fact, we were able to show that that many apparent counterexamples to Tracking Theories founder by overlooking this important feature of such theories (as made explicit by Dretske<sup>14</sup> and Nozick.<sup>15</sup> So it is a significant error to take Nozick's account of tracking to be only about beliefs and not about how one arrives at those beliefs. By proudly announcing his intention to focus only upon beliefs, Williams guarantees non-success at responding to our reply to his Backward Clock Case. In this context, Williams' claim that we are defending a different theory than Nozick's because we talk about **METHOD** fails. Rather, if one overlooks the crucial role that **METHODS** play in Nozick's account of knowledge (or **REASONS** in Dretske's account of knowledge) then one really just does not understand Nozick, Dretske or Tracking Theories. We suspect, in Williams' case, it is all three. This brings us to Dretske's Conclusive Reasons Account of Knowledge.

### 3. Dretske's Conclusive Reasons Account of Knowledge

Dretske's early Conclusive Reasons Account of Knowledge is, for many of us, his definitive account of knowledge. The account of knowledge contained in his book, *Knowledge and the Flow of Information*, provides a cognitive science, information-theoretic gloss that essentially reproduces the conclusive reasons story about knowledge. Without pretending to defend that claim, let's rehearse the account of knowledge contained in Dretske's Ur-text, "i.e., Conclusive Reasons." As Dretske states it:

S knows that *p* just in case

- (1) S believes that *p* (without doubt, reservation or question) on the basis of R.
- (2) R would not be the case unless *p* were the case.
- (3) Either S knows that R, or R is some experiential state of S.<sup>16</sup>

The subjunctive condition, 2, is to be read as saying that: Given R (your reasons or evidence), and fixed circumstances C (all those conditions that are logically and causally independent of the fact that *p*), then it is not physically possible that not *p*. Williams makes his first mistake here by construing Dretske's notion of a reason as referring only to a premise in an argument from an

---

<sup>14</sup> Fred Dretske, "Conclusive Reasons," *Australasian Journal of Philosophy* 49, I (1971): 1-22.

<sup>15</sup> Nozick, *Philosophical Explanations*.

<sup>16</sup> Dretske, "Conclusive Reasons," 12-13.

internalist perspective. Dretske does not restrict the idea of a reason in that way, but includes one's evidence, i.e., facts one knows to obtain, even if one is not aware of what one's evidence is. Dretske is, after all, an externalist about knowledge. Referring to Dretske's account here, Williams misunderstands Dretske in claiming that: "Here R is a reason that S has for believing that *p*. We nowhere talked of a reason."<sup>17</sup> At any rate, Williams goes on to argue that Dretske's sensitivity condition is satisfied because of the conjunctive reason that the

...hands point to 4:30 p.m. and your clock has always worked perfectly reliably. But this conjunction would not be true unless it were 4:30 p.m., because the hands would not point to 4:30 unless it were 4:30 p.m. This is because the circumstances in which you find yourself include those in which the clock runs perfectly reliably backwards from 5:00 p.m. to 4:00 p.m. Finally, we may stipulate that you know the conjunction that the hands point to 4:30 p.m. and your clock has always worked perfectly reliably. (1) to (3) are all true, but you do not know that it is 4:30 p.m. any more than you know this in Stopped Clock. So Dretske's early analysis is also too weak, predicting knowledge where there is ignorance.<sup>18</sup>

In this case, the key question for Dretske would be whether given your reasons or evidence R, and fixed circumstances C, it is physically possible that not *p*. That is, given that the clock displays 4:30 p.m. and has always been reliable, and the fixed circumstances surrounding the production of that clock, is it physically possible that it is not 4:30 p.m.? It is important to note here that Dretske intends the notion of what is physically possible to be constrained by natural law and the circumstances at hand. That is, could that clock have read 4:30 when it was not 4:30 p.m.? As we pointed out in our reply to Williams, the answer to this question for Dretske is "Yes, it is physically possible that the clock could have read 4:30 when it was not 4:30 p.m. at all." As we noted: "The clock in the Normal Clock case wouldn't have said that the time was 4:30 by displaying "4:30" if it hadn't been 4:30. Ted's clock, however, might have done this even if it hadn't be 4:30."<sup>19</sup> The cult might have devised the clock to read 4:30 when it was not 4:30 p.m. by making it run perfectly well backwards but more slowly such that it never recorded the correct time at any moment during that hour.<sup>20</sup> Hence, Dretske's sensitivity condition is not satisfied. You can learn things from people, Dretske said, but only from people who would not say something unless it were true. Applied to our case, this suggests that the cult clock-makers cannot be trusted

---

<sup>17</sup> Williams, "There's Nothing," 374.

<sup>18</sup> Williams, "There's Nothing," 375.

<sup>19</sup> Adams et.al., "Beat the (Backward)," 358.

<sup>20</sup> Adams et.al., "Beat the (Backward)," 357-358.

Murray Clarke, Fred Adams, and John A. Barker

because they made a clock that, in effect, lies for 59 of 60 minutes from 4:00 p.m. until 5:00 p.m. Anyone that makes a clock like that cannot be trusted, they might easily have made other deceptive clocks that would lie in all sorts of physically possible ways. But Williams responds to the suggestion that the cult clock makers might have made a clock that was slower by countering the final claim about Ted by us, saying:

This last claim is false. We stipulated that in the actual world, the clock runs perfectly reliably backwards from 5:00 p.m. to 4:00 p.m. So the only time at which its hands can point to 4:30 p.m. is when it is 4:30 p.m. Adams et.al.<sup>21</sup> point out that the cult could design the clock so its hands wouldn't point to the correct time at any time during the hour that you nap (say by making it run backward more slowly). Perhaps they had that possibility in mind. But as we described Backward Clock, worlds close to the actual circumstances in which you look at it cannot include those in which its mechanism differs from that which makes it run perfectly reliably backwards from 5 p.m. to 4:00 p.m. As we said, this is because the truth-adherence of your belief that it is 4:30 in Normal Clock resides in the fact that you would still have that belief in slightly changed circumstances in which the mechanism of the clock continues to work perfectly reliably. Likewise, the worlds close to the actual circumstances of Stopped Clock surely include those in which the mechanism of the clock is stopped.

What is essential to our counterexample then, is that the behaviour of its mechanism gets fixed across close possible worlds. Anything else, including the intentions of its designers, is simply irrelevant. In fact we introduced the story of the cult into the example to ensure that the behaviour of its mechanism gets fixed across close possible worlds, but other stories could be told. Perhaps the cult intended to symbolize the cursed nature of the hour with a seemingly unnatural phenomenon. Indeed we could dispense with the cult entirely and stipulate that a bug in the programming of the microchip circuit of your clock causes it to run perfectly reliably backwards from 5:00 p.m. to 4:00 p.m. during a particular hour.<sup>22</sup>

What is crucial then, for Williams, is that the mechanism of the clock is held fixed across close possible worlds when considering subjunctive conditionals of the sort that Nozick imposes on knowledge. Unfortunately for Williams and Sinhababu, what needs to be held fixed across possible worlds is not **mechanisms** but the **method M** for Nozick, or, for Dretske, the **circumstances C** relative to the reasons or evidence R. Methods are determined to be methods from the inside for Nozick. As he notes:

---

<sup>21</sup> Adams et.al., "Beat the (Backward)."

<sup>22</sup> Williams, "There's Nothing," 372.

A person can use a method (in my sense) without proceeding methodically, and without knowledge or awareness of what method he is using. Usually, a method will have a final upshot in experience on which the belief is based, such as visual experience, and then (a) no method without this upshot is the same method, and (b) any method experientially the same, the same 'from the inside,' will count as the same method. Basing our beliefs on experiences, you and I and the person floating in the tank are using, for these purposes, the same method.<sup>23</sup>

So, for instance, the Grandmother uses the method of visual perception in arriving at beliefs in the actual world and it is this method that must be held fixed in near possible worlds when considering whether the Grandmother knows that her Grandson is well in the actual world. Nowhere does Nozick claim that all circumstances must be held fixed, including circumstances that are logically and causally dependent on the fact expressed by *p*. Dretske, in fact, is explicit about this. The circumstances, *C*, that are held fixed when considering subjunctive conditionals relating *R* and *p* like his (2) are all those circumstances that are logically and causally independent of the state of affairs expressed by *p*.<sup>24</sup> As Dretske notes: "But does *C* include all the circumstances that prevail on the occasion in question or only some of these? Clearly not all the circumstances since this would trivialize every subjunctive conditional of this sort."<sup>25</sup> Dretske's idea here is that it cannot be the case that all circumstances are held fixed or all subjunctive conditionals would be trivially true. But that is not the case.

In the case of the Backward Clock, then, one cannot hold the mechanism of the clock fixed in near possible worlds because that circumstance is something that is dependent on the fact that *p*: that it is 4:30 p.m. That is, there is a dependency relationship between the fact that it is 4:30 p.m. and that the mechanism works the way that it does. Change the mechanism, and you change the time. We need to ask, therefore, "If it were not the case that the time was 4:30 is it physically possible that you would believe that it was 4:30 p.m. anyway?" The answer is Yes, the cult might have made the clock display 4:30 p.m. when it was not 4:30 p.m. because they might have made it lie in another way. If they can make it lie one way, then there are many ways that they can make the clock lie. If there is one way, then there are many ways. The point is that the clockmakers made the mechanism causally dependent on the time. The clock flips to 5:00 p.m. at 4:00 p.m. Hence, the clock's mechanism is causally dependent on *p*: "It is 4:30 p.m." As such, the clock's mechanism must be allowed to vary. You cannot hold the mechanism fixed as Williams and Sinhababu wish to do, and must insist on, to

---

<sup>23</sup> Nozick, *Philosophical Explanations*, 185.

<sup>24</sup> Dretske, "Conclusive Reasons."

<sup>25</sup> Dretske, "Conclusive Reasons," 48.

make their purported counterexample work, because it is causally dependent on the fact that the time was manipulated at 4:00 p.m. That mechanism must be allowed to vary or the relevant counterfactual conditional, i.e., The clock would not have read 4:30 p.m. unless it were 4:30 p.m., would be trivialized. As Dretske says concerning defective thermometers, in a parallel case, you cannot trust defective thermometers: "If it is that kind of thermometer, then if S's only basis for thinking his child's temperature normal is a 98.6 reading on it, then he does not know that his child's temperature is normal. It might be normal, of course, but if S knows that it is, he must have more to go on than the reading on this (defective) thermometer."<sup>26</sup> Likewise, you cannot trust defective clocks. If Williams and Sinhababu **could insist** on holding the mechanism **fixed** then they **could** have made their case. However, Dretske is explicit that this is not possible and Nozick's sensitivity condition imposes the same constraint here as Dretske's does: to allow the mechanism to be fixed across near possible worlds would be to **trivialize** Nozick's sensitivity condition, a condition that is equivalent to Dretske's sensitivity condition.

Another important point about method comes out in the claim, from the long quote from Williams a few pages back, that Williams and Sinhababu could have dispensed with talk of the cult entirely and just had a bug in the microchip of the circuit of your clock cause the clock to run perfectly reliably backwards from 5:00 p.m. to 4:00 p.m. Call this the 'Buggy Clock Case.' This buggy clock is no longer properly calibrated once the bug kicks in, and it becomes a 'broken' or 'improperly functioning' clock analogous to a stopped clock, analogous to a clock that happens to stop at noon one day and happens to restart at noon the next day, etc. The design aspect of such clocks is no longer operative, and therefore they don't really **SAY** anything about the time, even though they continue to **DISPLAY** the time and appear to **SAY** something about it. We want here to underline a crucial point about the nature of method or reasons for Nozick and Dretske. There is a world of difference between what a clock **DISPLAYS** and what it **SAYS**. The first is a pre-reflective matter, the second involves interpretation and method. The design aspect of clocks only becomes clear if they function in the way that they are supposed to, only if what they **SAY** accords with their design. Moreover, in fact, sometimes people have to learn to 'read' a clock, i.e., learn to interpret what the display says. Instructions from the designer (or manufacturer) will accompany a watch in such cases. In this respect, what Williams has to say about method is out of step with what externalists like Nozick and Dretske intend. Williams is not sensitive to the distinction between what the clock **DISPLAYS**, i.e., the position of

---

<sup>26</sup> Dretske, "Conclusive Reasons," 2.

its hands, and what the clock SAYS, i.e., what it designedly indicates about the current time, and how that distinction functions in Nozick's account of method and Dretske's account of reasons (in his broad notion of reason as including evidence). In fact, Williams evidently did not understand how we were using the term 'SAY' in our reply to him. The result is not only a misunderstanding of our view, but of Nozick's view and Dretske's view. It should also be noted that Williams talks about your **knowing** that the clock has always been reliable in the Backward Clock case, but that is the kind of **internalist** talk that externalists eschew. The possession of reasons or evidence, for Nozick and Dretske, does not require any sort of internal awareness or recognition in the epistemic internalist sense. One's having reasons or evidence is crucial to knowing but may involve no occurrent thought or access to a thought about the method at all. As Nozick says: "A person can use a method (in my sense) without proceeding methodically, and without knowledge or awareness of what method he is using."<sup>27</sup> All of this is compatible with the idea that what ultimately counts as the method is experiential states that are internal to the cognizer. In Nozick's sense of method: You and I, and a person floating in a tank on Alpha Centauri might be using the same method because methods are individuated from the standpoint of the cognizer. The cognizer, however, need not have any grasp of what that method is.

#### 4. Closing Remarks

The upshot of this is that Williams and Sinhababu have misread both Nozick and Dretske, since what is held fixed when considering Nozick's conditions 3 and 4 or Dretske's 2 is not mechanisms but the method *M* or reasons *R* in relation to *p* (such as that the Grandmother used the method of visual perception when determining the health condition of her Grandson or the person waking up and using visual perception to determine the time). Facts that are logically or causally independent of *p* are, however, held fixed. Put otherwise, facts that are dependent on *p* are allowed to vary. The mechanism of the clock is, therefore, dependent on the fact that *p*, and so that how that mechanism works must be allowed to vary or we would trivialize subjunctives regarding it. By holding the mechanism of the clock fixed in near possible worlds, Williams and Sinhababu have only succeeded in trivializing what are actually the profound implications of the employment of subjunctive conditionals in the articulation of two truth-tracking accounts of knowledge. Only by distorting the fundamental nature of Dretske's and Nozick's accounts of knowledge have Williams and Sinhababu provided the appearance of a

---

<sup>27</sup> Nozick, *Philosophical Explanations*, 185.

counterexample to truth-tracking accounts of knowledge. Hence, Williams cannot insist that the mechanism of the clock be held fixed in near possible worlds. We need to consider alternative near possible worlds where that mechanism is, for instance, simply slowed down and where we always get the time wrong to see that the observer could believe that it is 4:30 in such a world and be mistaken because all of the times are wrong. Such mechanisms, such signals, are much too equivocal to deliver knowledge in accordance with Nozick's or Dretske's accounts of knowledge. As such, they fail to track the truth in nearby possible worlds as promised. Dretske and Nozick tie beliefs via the Method M or Reasons R to facts. Those methods M or reasons R must be sensitive for both Dretske and Nozick and additionally adherent for Nozick in order to track the truth. The method M or Reasons R are held fixed but the circumstances that are logically and causally independent of the fact expressed by  $p$  are not held fixed on either account. This is the case because Dretske and Nozick wish to avoid the trivialization of the sensitivity subjunctive conditionals that they employ. The upshot is that Williams and Sinhababu have failed to advance a genuine counterexample to tracking-theories of knowledge.



# REPLY TO SIMION

Jonathan L. KVANVIG

**ABSTRACT:** Mona Simion questions whether there is a distinction between taking back an assertion and taking back only the content of an assertion, as I have claimed. After arguing against the distinction in question, Simion grants that there is a difference between the cases that I use to illustrate the distinction, and thus turns to the task of explaining the difference in a way that keeps it from undermining the knowledge norm. The explanation she offers is in terms of a distinction between doing something that is wrong and doing something that is blameworthy. I respond here by defending the distinction and questioning the explanation she gives of it.

**KEYWORDS:** knowledge norm of assertion, blameworthiness, normativity, assertion

Mona Simion<sup>1</sup> questions whether there is a distinction between taking back an assertion and taking back only the content of an assertion, as I claimed in arguing against the knowledge norm of assertion.<sup>2</sup> After arguing against the distinction in question, Simion grants that there is a difference between the cases that I use to illustrate the distinction, and thus turns to the task of explaining the difference in a way that keeps it from undermining the knowledge norm. The explanation she offers is in terms of a distinction between doing something that is wrong and doing something that is blameworthy.

I have elsewhere addressed the idea of salvaging the knowledge norm by appeal to a distinction between violating a norm and being blameworthy for doing so, both in “Norms of Assertion”<sup>3</sup> and more extensively in *Rationality and Reflection*,<sup>4</sup> especially chapters 2 and 3. The arguments there attack directly the idea that any distinction between blameworthiness and impropriety of some more fundamental sort, or more generally between any primary notion of propriety and some secondary notion, can explain away the purported counterexamples to the knowledge norm of assertion. I argue that such distinctions misunderstand the

---

<sup>1</sup> In Mona Simion, “Assertion: Just One Way to Take It Back,” *Logos & Episteme* VII, 3 (2016): 385-391.

<sup>2</sup> In Jonathan L. Kvanvig, “Knowledge, Assertion, and Lotteries,” in *Williamson on Knowledge*, eds. Duncan Pritchard and Patrick Greenough (Oxford: Oxford University Press, 2009), 140-160.

<sup>3</sup> Jonathan L. Kvanvig, “Norms of Assertion,” in *Assertion*, eds. Jessica Brown and Herman Cappellan (Oxford: Oxford University Press, 2011).

<sup>4</sup> Jonathan L. Kvanvig, *Rationality and Reflection* (Oxford: Oxford University Press, 2014).

nature of fundamental normativity, treating it in the way that is best reserved for some derivative domains, such as the legal sphere, where the normativity in question is partially a function of some more fundamental normativity.

Since the second part of Simion's paper does not engage with these arguments, I will bypass responding to that part of her paper here, since the general approach she takes is addressed already in the material cited above. I'll focus, then, on the claim that the distinction between two kinds of taking back is mistaken.

One example I used to illustrate the distinction is as follows:

For example, if we assert a claim and then are shown that the claim is false, we take back the content of our speech act, but we needn't apologize for or regret the very act itself. Randy says, "I've studied music all my life; there's no piece of group music even moderately well-known in the U.S. where part of the group is playing in 15/16 time and another part in 17/16 time," to which Michael responds, "That's certainly a reasonable judgment, except that you don't know enough about King Crimson. They are moderately well-known, and they have just such a piece." Michael then shows Randy the piece (so, I'm assuming that Michael is correct), to which Randy says, "I was wrong, I take it back." Now Randy may regret his assertion if he is the sort of person who strongly dislikes confronting his own fallibility. He may even vow to be much more careful not to say anything at all when he risks being wrong in order not to repeat this embarrassing moment, though such a response is surely overblown. Chagrin is normal, even mild embarrassment, but apologizing would be unctuous and overwrought. As I told the story, Randy responds appropriately. He doesn't apologize for making the assertion, but what he does instead is take back the content of the assertion. In fact, were he to apologize, the natural response would be dismissive: "Give it a rest, nobody's always right ..."<sup>5</sup>

Simion grants that there is a difference between the cases where an apology is appropriate and cases where it is not, but claims that the explanation of the difference can't be given in terms of a distinction between taking back the speech act itself and taking back its content:

But if the propositional content is inert in isolation, it is less clear how Kvanvig envisages one being able to take it back in isolation. To see this, notice that assertion, as opposed to other types of actions—say, having vacationed in Hawaii—can be 'taken back.' Not in the sense that one can change the past as to not have had asserted in the first place, of course. Rather, taking back an assertion that *p* refers to no longer standing behind the commitments implied by having asserted that *p*. Now, *p* itself, in isolation, does not imply any commitments whatsoever. That is, depending on which illocutionary force we

---

<sup>5</sup> Kvanvig, "Knowledge, Assertion," 148.

will act upon it with, different commitments will follow. If I promise that *p*, for instance, I commit myself to a future course of action; if I assert that *p*, I commit myself to, at least, it being the case that *p*.

If that is the case, it becomes clear that in order to take an assertion back, that is, to be released from the commitments implied by it, it has to be the case that I take back everything, force and content. I cannot only take back the content *p*, because *p* in isolation does not commit me to anything, inasmuch as I do not present it as true, or command *p*, or promise *p*, etc. Also, I cannot only take the action back either, because presenting nothing as true, or promising nothing also fails to imply any commitments on my part.<sup>6</sup>

Simion notes that content is inert in isolation, taking on various types of force depending on the kind of speech act in which the content is embedded. As a result, taking back the content of an assertion can't involve retracting some speech act itself, unless one takes back both the content and the assertion simultaneously. Hence, if the taking back is supposed to draw a distinction between taking back one kind of speech act versus taking back another kind of speech act, the distinction cannot be drawn.

It should be noted, however, that in the example I used above, as well as elsewhere in the paper, the distinction is not drawn in terms of two different kinds of speech acts. One side of the distinction applies to a speech act, for when one apologizes for, or regrets, an assertion, the object of one's attitude is the assertion itself, which is a speech act. But when one takes back only the content of an assertion, one does not have a speech act as the object of one's attitude nor of the act of taking back. Instead, the object of the taking back is whatever intellectual commitment to the claim led to the assertion in the first place, and what one is doing is countermanding that commitment. Thus, to take back the content of an assertion, as opposed to taking back the speech act itself, has as its object a commitment which is a mental state or act. In the usual case, such a commitment would be either a belief (a mental state) or the adoption of it (a mental act).

This distinction alone does not undermine the knowledge norm of assertion, but is merely one cog in a machine aimed at undermining that account. It is a defensive maneuver aimed at showing that the acknowledgement of a lapse of some sort, when it is pointed out that we don't know what we are talking about, is not the right kind of acknowledgement to justify endorsing the knowledge norm. So long as there are differences in this regard concerning the cases I describe, these cases can fulfill this defensive task whether or not the

---

<sup>6</sup> Simion, "Assertion," 287.

Jonathan L. Kvanvig

differences are properly characterized in terms of a distinction between taking back the assertion itself versus taking back its content. Simion objects to this way of explaining the differences, but, as I've argued, I don't think her concerns undermine this approach.

# UNDAUNTED EXPLANATIONISM

Kevin McCain

**ABSTRACT:** Explanationism is a plausible view of epistemic justification according to which justification is a matter of explanatory considerations. Despite its plausibility, explanationism is not without its critics. In a recent issue of this journal T. Ryan Byerly and Kraig Martin have charged that explanationism fails to provide necessary or sufficient conditions for epistemic justification. In this article I examine Byerly and Martin's arguments and explain where they go wrong.

**KEYWORDS:** epistemic justification, evidentialism, explanationism, explanationist evidentialism

Most everyone is an explanationist in the sense that they regularly employ explanatory reasoning/inference to the best explanation. This sort of reasoning is “so routine and automatic that it easily goes unnoticed.”<sup>1</sup> Recognition of the ubiquity of explanatory reasoning in both our everyday lives and our scientific practices helps make plausible a stronger sense of explanationism. Explanationists of this stronger variety not only accept inference to the best explanation as a legitimate form of reasoning, they contend that all epistemic justification is a matter of explanatory considerations. I am such an explanationist. Along with my fellow Alabama Explanationist, Ted Poston, I have worked to elucidate and defend explanationism in a number of works.<sup>2</sup> Fortunately, these efforts have helped spur on discussion and debate concerning the merits of explanationism.<sup>3</sup> Most recently,

---

<sup>1</sup> Igor, Douven (2011). “Abduction,” in *The Stanford encyclopedia of philosophy* (Spring 2011 Edition), ed. Edward N. Zalta, URL = <http://plato.stanford.edu/archives/spr2011/entries/abduction/>.

<sup>2</sup> Kevin McCain, “Explanationist Evidentialism,” *Episteme* 10 (2013): 299-315, “Evidentialism, Explanationism, and Beliefs about the Future,” *Erkenntnis* 79 (2014): 99-109, *Evidentialism and Epistemic Justification* (New York: Routledge, 2014), “Explanationism: Defended on All Sides,” *Logos & Episteme* 6, 3 (2015): 333-349, *The Nature of Scientific Knowledge: An Explanatory Approach* (Switzerland: Springer, 2016), “Explanationist Aid for Phenomenal Conservatism,” *Synthese* (forthcoming), Kevin McCain and Ted Poston, “Why Explanatoriness Is Evidentially Relevant,” *Thought* 3 (2014): 145-153, and Ted Poston, *Reason & Explanation: A Defense of Explanatory Coherentism* (New York: Palgrave-Macmillan, 2014).

<sup>3</sup> See, for example, T. Ryan Byerly, “Explanationism and Justified Beliefs About the Future,” *Erkenntnis* 78 (2013): 229-243, T. Ryan Byerly and Kraig Martin, “Problems for Explanationism on Both Sides,” *Erkenntnis* 80 (2014): 773-791, T. Ryan Byerly and Kraig Martin, “Explanationism, Super-Explanationism, Ecclectic Explanationism: Persistent Problems on Both

this discussion has been further advanced by T. Ryan Byerly and Kraig Martin's (hereafter "B&M") critique of explanationism in this journal.<sup>4</sup> In their article B&M contend that my recent attempt to defend explanationism fails.<sup>5</sup> They maintain that explanationism provides neither necessary nor sufficient conditions for epistemic justification. They are mistaken on both accounts. Nevertheless, B&M's objections are instructive to consider and respond to because doing so helps to clarify explanationism. This clarity helps demonstrate explanationism's viability as a theory of epistemic justification.

In this article my primary focus is B&M's most recent attack on explanationism. In light of this, I will not provide a general defense or motivation for explanationism. Instead, I will briefly explain explanationism and then turn to B&M's objections to it. I will explicate the problems with B&M's objections to both the sufficiency and necessity of explanationism. The end result is that despite interesting challenges from B&M explanationists should remain undaunted.

## 1. Explanationism

Before considering the objections that B&M press for explanationism it will be helpful to have a clear statement of the view. The central idea of explanationism is that epistemic justification is fundamentally a matter of explanatory relations. There are a variety of ways that one might spell out this key idea though. Hence, explanationism is perhaps best understood as more of a general approach, or a family of theories, than a particular theory of epistemic justification. Despite this fact, it will be useful to rely on the following specific explanationist theory in the present discussion for two reasons. First, this is a plausible explanationist theory of justification, and it works well as a test for objections to explanationism in general. Second, this is the specific explanationist theory that B&M target in their attack.<sup>6</sup> So, from this point on, unless otherwise noted, 'explanationism' will refer to the following account of justification:

---

Sides," *Logos & Episteme* 7 (2016): 201-203, William Roche and Elliott Sober, "Explanatoriness is Evidentially Irrelevant, or Inference to the Best Explanation Meets Bayesian Confirmation Theory," *Analysis* 73 (2013): 659-668, William Roche and Elliott Sober, "Explanatoriness and Evidence: A Reply to McCain and Poston," *Thought* 3 (2014): 193-199, Gregory Stoutenberg, "Best Explanationism and Justification for Beliefs about the Future," *Episteme* 12 (2015): 429-437, and Gregory Stoutenberg and Bryan Appley, "Two New Objections to Explanationism," *Synthese* (forthcoming).

<sup>4</sup> Byerly and Martin, "Explanationism."

<sup>5</sup> McCain, "Explanationism: Defended."

<sup>6</sup> They refer to this as 'super-explanationism.'

A person, S, with evidence *e* at time *t* is justified in believing *p* at *t* if and only if at *t* S

has considered *p*, and:

- (i) *p* is part of the best explanation available to S at *t* for why S has *e*  
or
- (ii) *p* is available to S as an explanatory consequence of the best explanation available to S at *t* for why S has *e*.<sup>7,8</sup>

There are a number of points about explanationism that could stand further clarification, and there are many things that can be said in support of this principle.<sup>9</sup> Here, however, it will be sufficient to expound upon just three points related to explanationism. The first point is that S's evidence, *e*, should be understood to be S's total evidence, not merely a proper subset of S's evidence. This is a key point that I have emphasized repeatedly in my various defenses of explanationism.<sup>10</sup> Focusing on a portion of S's evidence rather than S's total evidence ignores the potential impact of defeating evidence. Failure to consider S's total evidence when applying explanationism can lead to misguided objections to the theory as will become clear in the next section.

The second point that needs clarified is what it is that makes an explanation the best. A variety of explanatory virtues contribute to making one explanation better than another. Things like "simplicity, explanatory power (the range of phenomena explained and/or how illuminating the explanation is), consistency with [background information], non-ad hocness, [and] predictive power (making

---

<sup>7</sup> McCain, "Explanationism: Defended," 339. In that article I refer to this account as "Ex-EJ 2.0". This account of propositional justification is a key component of my complete theory of epistemic justification, *Explanationist Evidentialism*, which is a theory of doxastic justification/well-founded belief. *Explanationist Evidentialism* is explained more fully in my book, *Evidentialism and Epistemic Justification*.

<sup>8</sup> Two important qualifications of explanationism are made clear in footnote 19 of "Explanationism: Defended". "In order for S to be justified in believing that *p* it must not only be the best available explanation of S's evidence, it must also be a sufficiently good explanation of S's evidence. Similarly, in order for S to be justified in believing an explanatory consequence, *p*, of the best available explanation of her evidence it has to be that the best available explanation of her evidence would explain *p* significantly better than it would  $\sim p$ ." Both B&M's criticism of explanationism and the current discussion assume that these conditions are satisfied in the examples in question.

<sup>9</sup> See McCain, "Explanationism: Defended" and *Nature of Scientific Knowledge* for discussion and defense of this exact formulation of explanationism. See McCain, "Explanationist Evidentialism", "Beliefs about the Future", and *Evidentialism* for defense of a similar principle that appeals to logical entailment rather than explanatory consequence.

<sup>10</sup> Poston, *Reason & Explanation*, makes this point explicit as well.

novel predictions)” are commonly appealed to as explanatory virtues.<sup>11</sup> Just as with inferences to the best explanation, the best explanation referred to in explanationism is the one that has the most favorable mixture of these virtues. In some cases it is easy to determine the best explanation. For example, if two potential explanations are alike in all ways except that one is consistent with background information and the other is not, then the former is better than the latter. In other cases it is not easy to determine which explanation is best. When one explanation is simpler than another, but the more complex explanation has more explanatory power it can be difficult to determine which is the best. Fortunately, for present purposes it is not necessary to provide precise details for determining when one explanation is better than another. Instead, it is enough that it is recognized that there are a variety of explanatory virtues and that the best explanation is the one with the most favorable balance of these virtues. It is a mistake to fixate on one explanatory virtue to the exclusion of the others. This point will also be further illustrated in the discussion of the next section.

Finally, the third point that needs to be clarified is what it means for  $p$  to be an explanatory consequence of the best explanation available to  $S$  at  $t$ . The idea here is fairly straightforward.  $p$  is an explanatory consequence of the best explanation available to  $S$  when “ $p$  would be better explained by the best explanation of  $S$ ’s evidence available to  $S$  at  $t$  than  $\sim p$  would. In other words, if  $p$  were true, the best available explanation of  $S$ ’s evidence would better explain its truth than it would the truth of  $\sim p$ , if  $\sim p$  were true.”<sup>12</sup> With these points in hand it is time to turn toward the substance of B&M’s attack on explanationism.

## 2. The Attack on Sufficiency

B&M begin their most recent critique of explanationism by arguing that it fails to provide sufficient conditions for justification. More specifically, they argue that there are cases where  $p$  is part of the best available explanation for why  $S$  has  $e$ , and yet  $p$  is not justified for  $S$ . It is worth quoting the example that they rest their case for this point on in its entirety:

Imagine that Sally is the lead detective on an investigation of a burglary. She typically uses an eight-step investigative procedure for crimes of this sort and this procedure involves gathering and analyzing multiple kinds of evidence – physical evidences, forensic evidences, testimonial evidences, psychological evidences, circumstantial evidences, and so on. Sally is now mid-way through her investigative procedure, having completed four of the eight steps. She has

---

<sup>11</sup> McCain, *Nature of Scientific Knowledge*, 159.

<sup>12</sup> McCain, “Explanationism: Defended,” 339.



gathered and analyzed the appropriate evidence for these four steps, but has not yet gathered or analyzed evidence that may or may not arise during the final four steps. The list of suspects with which Sally began has been narrowed, and there is one very promising suspect in particular named Jeremy. In fact, the claim (call this the Jeremy hypothesis) is the best explanation available to Sally for all of the evidence she currently has obtained through the first four steps. There are multiple witnesses locating someone who fits Jeremy's description at the scene of the crime at the time at which it was committed. Some drug paraphernalia like that which Jeremy commonly uses to feed his drug habit was found at the scene of the crime. Jeremy seems to display a sense of satisfaction or gladness about the robbery. His bank account reflects a deposit shortly after the incident. Other current suspects, while not ruled out, do not fit the evidence Sally currently has anywhere nearly as well as Jeremy does. The Jeremy hypothesis is the best available explanation for the evidence Sally currently has and it is a very good explanation of that evidence.

But Sally isn't justified in believing the Jeremy hypothesis. For, she has good reason to think that there may very well be relevant evidence concerning the burglary that she does not currently have. After all, there have been many times in the past where, after completing step four of her investigation, things took a dramatic swing. It has not at all been uncommon that at these later stages in the process, an alternative suspect emerges who fits the data even better than previous suspects. Thus, while the Jeremy hypothesis is the best available explanation of the evidence Sally currently has, and while it is even a very good explanation of that evidence, Sally is not justified in believing this hypothesis. Believing the Jeremy hypothesis would be premature. The correct explanation for Sally's data may very well not be available at present, and she has good reason to think this.<sup>13</sup>

B&M claim that in this case the Jeremy hypothesis is the best explanation of Sally's evidence, but she is not justified in believing it. In light of this they claim that satisfying the conditions of explanationism is not sufficient for having justification.

In my previous defense of explanationism I pointed out that it seems B&M overlook the requirement that in order to truly satisfy explanationism the proposition (or hypothesis) in question must be part of the best explanation of S's *total evidence*.<sup>14</sup> Importantly, I noted that the best explanation of one's total evidence does not have to be a specific hypothesis; it can be general. The upshot of this earlier discussion is that the explanationist can readily agree that the Jeremy hypothesis is part of the best explanation of a proper subset of Sally's evidence, but this does not pose a problem for explanationism because a more general

---

<sup>13</sup> Byerly and Martin, "Problems for Explanationism," 783.

<sup>14</sup> McCain, "Explanationism: Defended."

explanation is the best explanation of her total evidence. Hence, the Jeremy hypothesis does not satisfy the conditions of explanationism, and the fact that Sally is not justified in believing this hypothesis poses no problem for explanationism.

Despite my earlier attempt to assuage their worries, B&M remain unconvinced. They say, “The problem with McCain’s response is easy to spot...McCain has simply overlooked what we said.”<sup>15</sup> B&M maintain that the Jeremy hypothesis is the best explanation of Sally’s evidence – it is superior to the sort of general hypothesis that I proposed. In support of this they point out that whereas a more general hypothesis such as <Some, as yet unknown, suspect committed the burglary> “does not predict all of the relevant data in the example...the Jeremy hypothesis does.”<sup>16</sup> B&M point out that the more general hypothesis does not predict “Jeremy’s attitude, the facts about his bank account, the reports of eyewitnesses of someone fitting Jeremy’s description, or the presence of drug paraphernalia of the same kind known to be employed by Jeremy,” but, of course, the Jeremy hypothesis does.<sup>17</sup> Consequently, B&M conclude that my attempt to defend explanationism fails because the Jeremy hypothesis really is the best explanation of Sally’s evidence, but she is not justified in believing that it is true.

Although B&M are certainly correct to emphasize that making accurate predictions is an important explanatory virtue, it seems that they may be fixating on this virtue to the exclusion of others. Additionally, they still seem to be failing to appreciate the qualification that explanationism is restricted to total evidence. In order to see these facts it will be helpful to flesh out some the details of B&M’s case and make things clearer.

Recall, B&M tell us that Sally has an eight-step process that she goes through when determining who committed a particular crime. Presumably, her process is very accurate when all eight steps are completed. At this point in the process, Sally has completed the first four steps and determined that Jeremy is by far the most likely suspect. B&M make it clear that while there are other suspects, which have not been conclusively ruled out, the claim that any one of these suspects committed the crime is an inferior explanation when compared with the claim that Jeremy committed the crime (the Jeremy hypothesis). They add to this that Sally has good reason to think that there is relevant evidence about the burglary that she currently lacks, “many times in the past” things change in the

---

<sup>15</sup> Byerly and Martin, “Explanationism,” 204.

<sup>16</sup> Byerly and Martin, “Explanationism,” 204.

<sup>17</sup> Byerly and Martin, “Explanationism,” 204.

later stages of her process, and “it has not at all been uncommon that at these later stages in the process, an alternative suspect emerges who fits the data even better than previous suspects.” Given all of this information, it seems that Sally has two sets of evidence. On the one hand, she has E (her current evidence pertaining to the burglary from steps one through four). On the other hand, she has E\* (evidence about how her investigative process works). B&M insist that explanationism has a problem because the Jeremy hypothesis is the best explanation of E – it predicts the relevant data and offers a better explanation than all its available rivals. Yet, they also maintain that Sally is not justified in believing the Jeremy hypothesis because of E\*. What should an explanationist say in response?

The correct response here is simple. B&M’s case relies on ignoring the qualification that explanationism should be understood in terms of total evidence rather than just a portion of the evidence. The Jeremy hypothesis is the best explanation of E, but it is not the best explanation of Sally’s total evidence, which includes both E and E\*. The reason for this is that Sally’s total evidence includes strong inductive evidence (E\*) for thinking that the hypothesis that is the best explanation at step four in the process is not true. That is, Sally’s total evidence includes inductive evidence for thinking that the Jeremy hypothesis is not true. As B&M describe this case Sally’s inductive reasons are strong enough to make it so that she should not believe the Jeremy hypothesis. If this is so, then it is plausible that the Jeremy hypothesis is not the best explanation of Sally’s total evidence. Instead, Sally’s total evidence is better, or at least equally well, explained by the more general hypothesis <Some, as yet unknown, suspect committed the burglary>. After all, B&M tell us that “it has not at all been uncommon that at these later stages in the process, an alternative suspect emerges who fits the data even better than previous suspects.”

Of course, one might complain, as B&M do, that the explanationist cannot plausibly maintain that the general hypothesis <Some, as yet unknown, suspect committed the burglary> is a better explanation than the Jeremy hypothesis. The Jeremy hypothesis explains relevant data that the general hypothesis does not – all sorts of facts about Jeremy, his bank account, and so on. B&M go so far as to claim that the Jeremy hypothesis predicts “*all* of the relevant data.”<sup>18</sup> Consequently, one might maintain that while it is true that Sally should not believe the Jeremy hypothesis, explanationists cannot plausibly claim this because the Jeremy hypothesis is the best explanation of Sally’s evidence.

---

<sup>18</sup> Byerly and Martin, “Explanationism,” 204. My emphasis.

There are two problems with this complaint. First, contrary to what B&M claim, the Jeremy hypothesis does not predict *all* of the relevant data. It is true that the Jeremy hypothesis predicts much (or perhaps all) of E, but this is not all of the relevant data! Sally's total evidence includes E and E\*. The Jeremy hypothesis does not predict E\*. This evidence is better explained by the general hypothesis. So, the fact that the Jeremy hypothesis predicts a portion of Sally's evidence should not be taken to be conclusive evidence for thinking that it is the best explanation of her total evidence.

Second, while predicting relevant data is an important explanatory virtue, it is not the only virtue. Fixating on the Jeremy hypothesis' ability to predict a portion of Sally's evidence is a mistake. There are many other explanatory virtues. An important explanatory virtue that bears particularly on this case is conservatism, or fit with background information. Although the Jeremy hypothesis is an excellent explanation when we consider only E, it is not a very good explanation when we consider Sally's total evidence because it fails to fit with her background information about how her investigative process tends to turn out. To see how background evidence can affect the quality of an explanation, consider a situation where you see a particular object. This object looks like a duck, quacks like a duck, walks like a duck, and so on. The best explanation of this information taken in isolation is that the object you see is a duck. However, this is not the best explanation when you add more details such as that you know that you are near a factory that produces large quantities of robotic duck decoys that are very hard to distinguish from real ducks, you are currently in the middle of the desert where ducks cannot survive, and so on. Although the 'duck hypothesis' predicts a lot of your evidence, its failure to fit with background evidence makes it a poor explanation overall. Similar considerations apply to B&M's Jeremy hypothesis. There are good grounds for denying that the Jeremy hypothesis really is the best explanation of Sally's *total* evidence. As a result, explanationism is not threatened by B&M's attack on its sufficiency.

### 3. The Attack on Necessity

Much of the recent debate concerning whether explanationism provides a necessary condition for justification has focused on the justification we have for beliefs about the future.<sup>19</sup> B&M seem to accept that while earlier incarnations of explanationism did not adequately account for the justification that we have for

---

<sup>19</sup> See Byerly, "Beliefs about the Future," Byerly and Martin, "Problems for Explanationism," McCain, "Beliefs about the Future," McCain, *Evidentialism*, and McCain, "Explanationism: Defended."

believing things about the future, the formulation of explanationism above does.<sup>20</sup> Nevertheless, they claim that explanationism still faces problems when it comes to providing a necessary condition for justification. Specifically, B&M argue that cases involving mathematical/logical entailment pose an insuperable problem for explanationism.

B&M's most recent discussion of cases of mathematical/logical entailment is very helpful.<sup>21</sup> In fact, their discussion has led me to recognize that my previous treatment of this sort of case needs revision.<sup>22</sup> Previously, I attempted to account for the justification that we have for beliefs concerning the conclusion of inferences involving appeals to mathematical/logical entailment solely in terms of explanationism's first condition, (i). B&M, however, have helped me see that this is not the best way to understand mathematical/logical entailments when they are combined with empirical claims.

Although B&M are correct in claiming that my earlier discussion of these cases in this journal is problematic, they are mistaken in thinking that explanationism lacks the resources to adequately accommodate cases of mathematical/logical entailment. Rather than simply discuss this issue in the abstract or go back to discussing cases that I have dealt with before, I will explain how explanationism does provide the correct results in this sort of case by examining B&M's latest example – a case that they claim poses a problem for all explanationist views, not just explanationism as formulated here.

B&M make use of 'surprising correlations' to press their objection to the necessity of explanationism. Here is their case:

Most years between 1999 and 2009 where Nicholas Cage appeared in at least 2 films were years between 1999 and 2009 where there were at least 98 drownings, and most years between 1999 and 2009 where there were at least 98 drownings were years between 1999 and 2009 where Cage appeared in at least 2 films. Now, imagine that someone, Joe, comes to know this fact, but does so without coming to know the number of Cage films and drownings for any particular year. Suppose next that Joe learns that in some particular year in the interval, say 2006, Cage was in at least 2 movies. Depending upon exactly the strength of the correlation and the appropriate threshold for justification, it is plausible that Joe would be justified in believing that in 2006 there were at least 98 drownings.<sup>23</sup>

In my previous article in this journal I offered a way of responding to similar cases that would suggest the following way of accounting for Joe's

---

<sup>20</sup> Byerly and Martin, "Explanationism."

<sup>21</sup> Byerly and Martin, "Explanationism."

<sup>22</sup> McCain, "Explanationism: Defended."

<sup>23</sup> Byerly and Martin, "Explanationism," 211.

justification.<sup>24</sup> According to my earlier suggestion, Joe is justified in believing <There were at least 98 drownings in 2006> because this proposition is part of the best available explanation of his evidence. The reason I offered for this is that the truth of <There were at least 98 drownings in 2006> is part of the best available explanation for why Joe is aware, or has a seeming, that <There were at least 98 drownings in 2006> follows from items of his evidence – namely, <most years between 1999 and 2009 where there were at least 98 drownings were years where Cage appeared in at least 2 films> and <In 2006 Cage appeared in at least 2 films>. As B&M point out, though, this is not quite right. The best explanation of Joe's awareness/seeming is not that <There were at least 98 drownings in 2006> is true. As they aptly note, Joe's seeming that <There were at least 98 drownings in 2006> follows from his evidence "isn't explained by there being 98 drownings in 2006. Indeed, the seeming would persist even if there were no Cage films or drownings. It is just a seeming about what makes what probable, not about what there is in the world."<sup>25</sup> So, my previous response to this sort of case will not work.

Fortunately, the problem with my earlier response is not the result of any failing of explanationism. Explanationism has the means of providing the correct result in these cases in a plausible manner. Essentially, the problem is simply that my previous response cut out important steps. Here is what an explanationist should say about B&M's case. Part of the best available explanation of Joe's evidence is that, as noted above, <There were at least 98 drownings in 2006> follows from <most years between 1999 and 2009 where there were at least 98 drownings were years where Cage appeared in at least 2 films> and <In 2006 Cage appeared in at least 2 films>. It is also part of the best available explanation of Joe's evidence that <most years between 1999 and 2009 where there were at least 98 drownings were years where Cage appeared in at least 2 films> is true. After all, B&M tell us that Joe *knows* this fact to be true. Additionally, <In 2006 Cage appeared in at least 2 films> is part of the best explanation of Joe's evidence too because B&M inform us that Joe has learned this fact as well. So, Joe has as part of the best explanation of his evidence that <most years between 1999 and 2009 where there were at least 98 drownings were years where Cage appeared in at least 2 films> is true, <In 2006 Cage appeared in at least 2 films> is true, and <There were at least 98 drownings in 2006> follows from these propositions. It is exceedingly plausible that <There were at least 98 drownings in 2006> is true is available to Joe as an explanatory consequence of the best explanation available to him for his evidence. The truth of <There were at least 98 drownings in 2006> is

---

<sup>24</sup> McCain, "Explanationism: Defended."

<sup>25</sup> Byerly and Martin, "Explanationism," 211.

much better explained by the best explanation of Joe's evidence than its denial is. The best explanation of Joe's evidence provides a very good explanation of <There were at least 98 drownings in 2006>, but a very poor explanation of its denial.<sup>26</sup> In light of this, according to the second condition, (ii), of explanationism, Joe is justified in believing <There were at least 98 drownings in 2006>. The problem with my earlier response was that it focused solely on (i) of explanationism for handling this sort of case, but the full story makes use of (i) and (ii). This is the key insight that B&M's discussion illuminates. Nonetheless, explanationism yields the intuitively correct result.

What is the upshot here? B&M's arguments are helpful in that they move the discussion of explanationism forward by making it clear that my earlier response fails to express how explanationism should be understood to handle cases involving mathematical/logical entailments when they are combined with empirical claims. However, at the end of the day explanationism has the tools necessary for providing the correct result in this sort of case. So, B&M's attack on the necessity of explanationism, like the attack on sufficiency, fails. Thus, despite B&M's assertion that explanationism faces 'persistent problems on both sides,' explanationists should remain undaunted.<sup>27</sup>

---

<sup>26</sup> As I have noted in various works this is perhaps best accounted for by the commonly accepted fact that higher probabilities explain better than lower ones. See Michael Strevens, "Do Large Probabilities Explain Better?" *Philosophy of Science* 67 (2000): 366-390 for arguments in favor of this fact about explanations. Also, see Jonah Schupbach and Jan Sprenger, "The Logic of Explanatory Power," *Philosophy of Science* 78 (2011): 105-127 for defense of an account of explanatory power on which higher probabilities offer greater explanatory power.

<sup>27</sup> I am grateful to Matt Frise, Jon Matheson, and Ted Poston for helpful comments and discussion.





# WHY GETTIER CASES ARE STILL MISLEADING: A REPLY TO ATKINS

Moti MIZRAHI

**ABSTRACT:** In this paper, I respond to Philip Atkins' reply to my attempt to explain why Gettier cases (and Gettier-style cases) are misleading. I have argued that Gettier cases (and Gettier-style cases) are misleading because the candidates for knowledge in such cases contain ambiguous designators. Atkins denies that Gettier's original cases contain ambiguous designators and offers his intuition that the subjects in Gettier's original cases do not know. I argue that his reply amounts to mere intuition mongering and I explain why Gettier cases, even Atkins' revised version of Gettier's Case I, still contain ambiguous designators.

**KEYWORDS:** ambiguous designator, analysis of knowledge, Gettier cases, justified true belief, semantic reference, speaker's reference

## 1. Introduction

In a reply to my "Why Gettier Cases Are Misleading,"<sup>1</sup> Philip Atkins sets out to defend the "orthodox view in contemporary epistemology," according to which "Edmund Gettier *refuted* the JTB [Justified True Belief] analysis of knowledge" (emphasis added).<sup>2</sup> Before I address Atkins' objections against the argument I put forth in "Why Gettier Cases Are Misleading," I would like to point out a few things that I find rather peculiar about his reply. First, Atkins contends that "Gettier's two cases [...] are genuine counterexamples to the JTB analysis."<sup>3</sup> But then he proceeds to "*revise* Gettier's first case so that there is no such semantic failure [i.e., failure to refer to the semantic referent of 'coins']" (emphasis added). If Atkins needs to *revise* Gettier's Case I in response to my criticism against it, then that means that Gettier's original case is *not* a genuine counterexample to the JTB analysis. After all, if it were a genuine counterexample, then there would be no need to revise it; it would work against the JTB analysis just as it is. Of course, epistemologists have long recognized that Gettier's original cases are problematic. One problem with Gettier's original cases, which I discuss in the paper to which Atkins responds, is the problem of false lemmas. Many epistemologists have found

---

<sup>1</sup> Moti Mizrahi, "Why Gettier Cases Are Misleading," *Logos & Episteme* 7 (2016): 31-44.

<sup>2</sup> Philip Atkins, "Are Gettier Cases Misleading?" *Logos & Episteme* 7 (2016): 379-384.

<sup>3</sup> Atkins, "Are Gettier Cases," 379.

it problematic that the subjects in Gettier's original cases infer their candidates for knowledge from falsehoods, and so have constructed Gettier-style cases with "no false lemmas."<sup>4</sup> Curiously, Atkins does not mention any of this and proceeds to defend Gettier's original cases as if they are entirely unproblematic, even though he is aware of the fact that I discuss "several 'Gettier cases' besides the two that Gettier originated."<sup>5</sup> As a result, Atkins' paper presents a somewhat inaccurate picture of the state of the debate over the status of Gettier cases as a "refutation" of the JTB analysis of knowledge.

Speaking of "refutations," another thing I find rather peculiar about Atkins' reply to my "Why Gettier Cases Are Misleading" is his use of the term 'refutation' in conjunction with his hedging and *seeming* talk. On the one hand, Atkins claims that Gettier's original "cases *refute* the JTB analysis of knowledge" (emphasis added).<sup>6</sup> If such cases do indeed amount to a *refutation* of the JTB analysis of knowledge, however, it is difficult to see why Atkins needs to *hedge* his claims and engage in *seeming* talk. Here are a couple of examples from his paper:

*Many* have the *strong intuition* that Smith fails to know (I). [(I) The man who will get the job has ten coins in his pocket.] Since Smith is justified in believing (I), we *seem* to have a counterexample to the JTB analysis (emphasis added).<sup>7</sup>

*I cannot speak for everyone*, but *I* have the *strong intuition* that Smith fails to know (I\*). Since Smith is justified in believing (I\*), we *seem* to have a counterexample to the JTB analysis (emphasis added).<sup>8</sup>

Of course, Atkins is not doing something new here. Arguments from Gettier cases against JTB are nothing more than appeals to intuition. If these arguments are to count as *refutations* of the JTB analysis of knowledge, i.e., conclusive proofs that

---

<sup>4</sup> For the "no false lemmas" response to Gettier cases, see David M. Armstrong, *Belief, Truth, and Knowledge* (New York: Cambridge University Press, 1973), 152 and Michael Clark, "Knowledge and Grounds: A Comment on Mr. Gettier's paper," *Analysis* 24 (1963): 46-48. See also Robert K. Shope, *The Analysis of Knowing: A Decade of Research* (Princeton, NJ: Princeton University Press, 1983), 24 and the Appendix in John L. Pollock, *Contemporary Theories of Knowledge* (Totowa: Rowman and Littlefield, 1986). Cf. Michael Levin, "Gettier Cases Without False Lemmas," *Erkenntnis* 64 (2006): 381-392. An early so-called Gettier-style case without false lemmas can be found in Gilbert Harman, *Thought* (Princeton: Princeton University Press, 1973), 75. Cf. William G. Lycan, "On the Gettier Problem Problem," in *Epistemology Futures*, ed. S. Hetherington (New York: Oxford University Press, 2006), 148-168. Lycan defends JTB with the addition of the "no false lemmas" condition.

<sup>5</sup> Atkins, "Are Gettier Cases," 379.

<sup>6</sup> Atkins, "Are Gettier Cases," 382.

<sup>7</sup> Atkins, "Are Gettier Cases," 380.

<sup>8</sup> Atkins, "Are Gettier Cases," 381.

JTB is false, it must be the case that our intuitions about hypothetical cases, such as Gettier cases, perfectly track the epistemic facts about such cases. This assumption, however, is rather controversial,<sup>9</sup> especially in light of the empirical evidence from experimental philosophy and cognitive science.<sup>10</sup> So, again, by proceeding as if Gettier's original cases are entirely unproblematic, Atkins' paper presents a somewhat inaccurate picture of the state of the debate over the status of Gettier cases as a "refutation" of the JTB analysis of knowledge. Gettier's original cases are problematic in at least two respects. First, they involve inferences from falsehoods. Second, the arguments made on the basis of Gettier cases are appeals to intuition, which are themselves a rather controversial sort of arguments in philosophy.<sup>11</sup>

Finally, as the quotes above illustrate, Atkins insists that his intuition is that subjects in Gettier cases do not know that *p*. Clearly, since I have argued that Gettier cases are misleading, which means that we should not assign much, if any, evidential weight to the so-called "Gettier intuition," i.e., the intuition that *S* doesn't know that *p* in a Gettier case, I do not find Atkins' insistence that he shares the "Gettier intuition" to be compelling evidence against my argument to the effect that Gettier cases are misleading.

With these preliminary remarks in hand, I will now address Atkins' objections and his attempt to defend the claim that Gettier's original cases "are genuine counterexamples to the JTB analysis."<sup>12</sup>

## 2. Atkins' Defense of Gettier's Case I

Atkins aims to defend Gettier's Case I by modifying it such that it does not involve any ambiguous designators. In my "Why Gettier Cases Are Misleading," I argue that 'coins' in

(I) The man who will get the job has ten coins in his pocket<sup>13</sup>

is an ambiguous designator. Atkins offers a revised case in which there are no ambiguous designators, or so he claims.

---

<sup>9</sup> See Moti Mizrahi, "Don't Believe the Hype: Why Should Philosophical Theories Yield to Intuitions?" *Teorema: International Journal of Philosophy* 34 (2015): 141-158.

<sup>10</sup> See Moti Mizrahi, "Three Arguments Against the Expertise Defense," *Metaphilosophy* 46 (2015): 52-64.

<sup>11</sup> See Moti Mizrahi, "Does the Method of Cases Rest on a Mistake?" *Review of Philosophy and Psychology* 5 (2014): 183-197.

<sup>12</sup> Atkins, "Are Gettier Cases," 379.

<sup>13</sup> Edmund L. Gettier, "Is Justified True Belief Knowledge?" *Analysis* 23 (1963): 121-123.

Suppose that Smith has strong evidence for believing that Jones is the man who will get the job and that Jones is handsome. We can suppose that Smith is justified in believing that Jones is handsome based on seeing Jones in person. Smith makes a rudimentary logical inference and says the following:

(I\*) The man who will get the job is handsome.

It turns out that (I\*) is true, but not for the reasons that Smith thinks. For it turns out that Smith is the man who will get the job and that, unbeknownst to Smith, he is also handsome. I cannot speak for everyone, but I have the strong intuition that Smith fails to know (I\*). Since Smith is justified in believing (I\*), we seem to have a counterexample to the JTB analysis.<sup>14</sup>

At first, Atkins simply asserts that, as far as he can tell, “there is no semantic failure when Smith uses the predicate ‘is handsome’.”<sup>15</sup> But then he acknowledges that there is an ambiguous designator in this case after all. The ambiguous designator is ‘the man’. As Atkins himself writes, “The speaker’s referent [of ‘the man’] is Jones, whereas the semantic referent [of ‘the man’] is Smith himself.”<sup>16</sup>

Because of this, presumably, Atkins revises Gettier’s Case I for the second time, this time to remove the ambiguous designator ‘the man’.

Suppose again that Smith has strong evidence for believing that Jones is the man who will get the job and that Jones is handsome. Smith performs an existential generalization and says the following:

(I\*\*) There is someone who is both getting a job and handsome.

It turns out that (I\*\*) is true, but not for the reasons that Smith thinks. For it turns out that (I\*\*) is made true by Smith himself. Even though Smith is justified in believing (I\*\*), and even though (I\*\*) is true, I have the strong intuition that Smith fails to know (I\*\*).<sup>17</sup>

At this point, however, it looks like Atkins is simply engaging in intuition mongering.<sup>18</sup> He claims to have a “strong intuition that Smith fails to know (I\*\*),”<sup>19</sup> but offers no reasons whatsoever to think that Smith indeed does not know that (I\*\*) is the case. Perhaps Atkins has “the strong intuition that Smith fails to know (I\*\*)”<sup>20</sup> because he has been taught that that’s the “right” response to Gettier cases.

---

<sup>14</sup> Atkins, “Are Gettier Cases,” 380-381.

<sup>15</sup> Atkins, “Are Gettier Cases,” 381.

<sup>16</sup> Atkins, “Are Gettier Cases,” 382.

<sup>17</sup> Atkins, “Are Gettier Cases,” 382.

<sup>18</sup> See Moti Mizrahi, “Intuition Mongering,” *The Reasoner* 6 (2012): 169-170 and Moti Mizrahi, “More Intuition Mongering,” *The Reasoner* 7 (2013): 5-6.

<sup>19</sup> Atkins, “Are Gettier Cases,” 382.

<sup>20</sup> Atkins, “Are Gettier Cases,” 382.

Or perhaps Atkins has “the strong intuition that Smith fails to know (I\*\*)”<sup>21</sup> because Smith infers (I\*\*) from ‘Jones is the man who will get the job and Jones is handsome,’ which is false by stipulation. If so, then we run into the “no false lemmas” (or inference from falsehoods) problem again, which Atkins completely ignores in his reply to my “Why Gettier Cases Are Misleading,” as I have mentioned above.

More importantly, and again, as Atkins himself acknowledges, his second rendition of Gettier’s Case I still involves an ambiguous designator. For, as Atkins himself writes, “there is some sense in which Smith has Jones in mind when inferring (I\*\*).”<sup>22</sup> So, as Atkins himself admits, his second rendition of Gettier’s Case I is a case of reference failure after all. Atkins dismisses this referential ambiguity by simply asserting without argument that “this point *seems* irrelevant” (emphasis added).<sup>23</sup> It might *seem* irrelevant to Atkins, and Atkins offers no reasons to think that it is irrelevant, but it *isn’t* irrelevant. In fact, it is an objection I address in the paper to which Atkins is replying. As I argue in “Why Gettier Cases Are Misleading,” if the candidates for knowledge in Gettier cases contain ambiguous designators, then that means that the relevant beliefs are ambiguous between two interpretations: “an ‘objective’ interpretation in terms of the conditions that make the belief true (i.e., in terms of semantic reference or what a speaker’s words mean) and a ‘subjective’ interpretation in terms of what *S* means (i.e., in terms of speaker’s reference or what a speaker means in uttering certain words).”<sup>24</sup> In Atkins’ second rendition of Gettier’s Case I, then, the belief that there is someone who is both getting a job and handsome is ambiguous between these two interpretations:

*Objective interpretation* (semantic reference): the semantic referent of ‘someone’ in <there is someone who is both getting the job and handsome> is the actual person that makes <there is someone who is both getting the job and handsome> true; otherwise, <there is someone who is both getting the job and handsome> would not be true.

*Subjective interpretation* (speaker’s reference): the speaker’s referent of ‘someone’ in <there is someone who is both getting the job and handsome> is what Smith has in mind when he believes that there someone who is both getting the job and handsome, which is Jones, not Smith himself, who is actually the person that makes <there is someone who is both getting the job and handsome> true.

---

<sup>21</sup> Atkins, “Are Gettier Cases,” 382.

<sup>22</sup> Atkins, “Are Gettier Cases,” 383.

<sup>23</sup> Atkins, “Are Gettier Cases,” 383.

<sup>24</sup> Mizrahi, “Why Gettier Cases,” 43.

Interpreted “objectively,” or in terms of what the words mean, <there is someone who is both getting the job and handsome> is not what Smith actually believes, since Smith uses ‘there is someone who’ to talk about what his evidence leads him to believe, which is “that Jones is the man who will get the job and that Jones is handsome,”<sup>25</sup> not that Smith is the man who will get the job and that Smith is handsome. Interpreted “subjectively,” or in terms of what Smith means by uttering these words, <there is someone who is both getting the job and handsome> is strictly false, since Smith uses ‘there is someone who’ to talk about something that does not in fact fulfill the conditions for being the semantic referent of ‘someone’ in this case.

In other words, Smith’s belief that there is someone who is both getting a job and handsome is ambiguous between two interpretations:

1. *Semantic reference*: There is someone (= Smith) who is both getting the job and is handsome.
2. *Speaker’s reference*: There is someone (= Jones) who is both getting the job and is handsome.

By stipulation, (2) is false, since it turns out that Smith gets the job. On (2), then, Smith simply has a false belief. On the other hand, (1) is not actually what Smith believes in this case, since Smith wishes to talk about Jones, which is what Smith’s evidence is about. To put it crudely, on (1), what goes on in Smith’s head does not match the facts of the case. Given this ambiguity, then, Atkins’ second rendition of Gettier’s Case I, like Gettier cases in general, is misleading.

### 3. Atkins’ Defense of Gettier’s Case II

Atkins’ attempt to defend Gettier’s Case II looks like another instance of intuition mongering. Atkins simply recounts Gettier’s Case II, without revisions, and asserts that

*Many have the strong intuition that Smith fails to know (h). [(h) Either Jones owns a Ford or Brown is in Barcelona.] Since Smith is justified in believing (h), we seem to have a counterexample to the JTB analysis (emphasis added).<sup>26</sup>*

Atkins’ “strong intuition” notwithstanding, there is an ambiguous designator in Gettier’s Case II, as I point out in “Why Gettier Cases Are Misleading.” As Atkins himself writes, Smith’s evidence for (h) is that “Jones has at all times in the past owned a car, and always a Ford, and that Jones has just offered

---

<sup>25</sup> Atkins, “Are Gettier Cases,” 382.

<sup>26</sup> Atkins, “Are Gettier Cases,” 383.

Smith a ride while driving a Ford.”<sup>27</sup> As in Gettier’s Case I, then, there is a mismatch between what goes on in Smith’s head and the facts about the case. Another way to see this, in addition to the way I have described above, is the following. In Atkins’ second rendition of Gettier’s Case I, Smith reasons as follows:

a. Jones will get the job.

Therefore,

b. There is someone who will get the job.

c. Jones is handsome.

Therefore,

d. There is someone who is handsome.

Therefore,

e. There is someone who will get the job *and* there is someone who is handsome.

As we can see, Smith’s evidence supports (e), not the belief that the one who will get the job and the one who is handsome are *one and the same person*. To see why, note that the move from (a) to (b) and the move from (c) to (d) are instances of existential generalization. If Smith were to reason backwards, however, from (e) by existential instantiation, Smith could just as easily end up with the false belief that Jones will get the job instead of the true belief that Smith will get the job; hence the ambiguity in terms of the referent of ‘someone’; in Smith’s mind that someone is not Smith himself, but rather Jones, since that is what Smith’s evidence, i.e., (a) and (c), is about.

From a logical point of view, this counts as an instance of equivocation. According to Quine,

[t]he fallacy of equivocation arises [...] when the interpretation of an ambiguous expression is influenced in varying ways by immediate contexts [...], so that the expression undergoes changes of meaning within the limits of the argument.<sup>28</sup>

In Atkins’ revised version of Gettier’s Case I, Smith reasons from evidence about one thing (namely, Jones) to a conclusion that is made true by something else (namely, Smith). This switch in reference “within the limits of the argument” makes this case appear like a genuine counterexample to JTB, even though it is not.

Similarly, in Gettier’s Case II, Smith reasons as follows:

---

<sup>27</sup> Gettier, “Is Justified,” 122.

<sup>28</sup> W. V. Quine, *Methods of Logic*, 4<sup>th</sup> Ed (Cambridge: Harvard University Press, 1982), 56.

- i. "Jones has *at all times in the past within Smith's memory* owned a car" (emphasis added).<sup>29</sup>
- ii. "Jones has *at all times in the past within Smith's memory* owned a Ford" (emphasis added).<sup>30</sup>
- iii. "Jones has *just* offered Smith a ride while driving a Ford" (emphasis added).<sup>31</sup>

Therefore,

- iv. Jones owns a Ford.

Therefore,

- v. Either Jones owns a Ford or Brown is in Barcelona.

Contrary to what Atkins suggests, Smith cannot simply make "a rudimentary logical inference"<sup>32</sup> from (i)-(iii) to (v), since (v) does not follow from (i)-(iii). Rather, (i)-(iii) are evidence for (iv), and then Smith infers (v) from (iv) by "a rudimentary logical inference,"<sup>33</sup> namely, addition.

As stipulated, however, "Jones does *not* own a Ford, but is *at present* driving a rented car" (emphasis added).<sup>34</sup> Note the use of temporal terms, such as 'at all times in the past', 'just', and 'at present', which is crucial here. For Smith wishes to talk about the person who "has *at all times in the past within Smith's memory* owned a Ford" (emphasis added).<sup>35</sup> It just so happens that this person does not own a Ford *at present*. Of course, this sort of thing happens all the time; something could be true about a person at one point in time and then stop being true at a later point in time. The proposition 'George W. Bush is the President of the United States' was true from 2001 until 2009, but it was not true before 2001 and it is not true *at present*. The proposition 'Barack Obama is the President of the United States' is true now, but it will no longer be true after January 20, 2017. Suppose, then, that on January 21, 2017, an eight-year-old reasons as follows:

Barack Obama has at all times in the past within my memory been the US President.

Therefore,

---

<sup>29</sup> Gettier, "Is Justified," 122.

<sup>30</sup> Gettier, "Is Justified," 122.

<sup>31</sup> Gettier, "Is Justified," 122.

<sup>32</sup> Atkins, "Are Gettier Cases," 383.

<sup>33</sup> Atkins, "Are Gettier Cases," 383.

<sup>34</sup> Gettier, "Is Justified," 123.

<sup>35</sup> Gettier, "Is Justified," 122.



Barack Obama is the US President *at present* (where the present time is January 21, 2017).

The eight-year-old's belief that Barack Obama is the US President at present will be false on January 21, 2017. The problem is that 'Barack Obama' is referentially ambiguous in this context. The reference of 'Barack Obama' in 'Barack Obama has *at all times in the past within my memory* been the US President' was fixed at some particular time in the past, since this piece of evidence comes from memory, whereas 'Barack Obama' in 'Barack Obama is the US President *at present*' is supposed to pick out the *present* US President.

Similarly, Smith infers (iv) from evidence that is time-indexed to a particular time in the *past*, since Smith wishes to talk about the person who "has *at all times in the past within Smith's memory* owned a Ford" (emphasis added),<sup>36</sup> but that no longer pertains to the *present* time, since Jones "is *at present* driving a rented car" (emphasis added).<sup>37</sup> Accordingly, there is "an unsignaled shift in meaning"<sup>38</sup> in Smith's reasoning from "In the *past*, (i), (ii), and (iii) were the case" to "At *present*, (iv) is the case." For this reason, there is an ambiguity in Gettier's Case II. Unlike Gettier's Case I (and Atkins' revised versions of the case), however, the ambiguity is not in terms of the referent of 'someone' (i.e., Smith or Jones), but rather in terms of the time to which the relevant propositions are indexed (i.e., past or present).

Accordingly, Smith's evidence, i.e., (i)-(iii), supports the belief that Jones owns a Ford at  $t_1$  (in the past), not the belief that Jones owns a Ford at  $t_2$  (at the present time). To see why, note that, the move from (iv) to (v) is an instance of disjunction introduction. If Smith were to reason backwards, however, from (v) by elimination, Smith could just as easily end up with the false belief that Jones owns a Ford at  $t_2$  (at the present time) instead of the true belief that Jones owns a Ford at  $t_1$  (in the past);<sup>39</sup> hence the ambiguity in terms of the time to which the relevant belief is indexed; in Smith's mind his belief is indexed to a time in the past, not the present, since that is what Smith's evidence, namely, (i)-(iii), is about.

As with Gettier's Case I, from the point of view of argumentation theory, this counts as an instance of the fallacy of equivocation. According to Johnson and

---

<sup>36</sup> Gettier, "Is Justified," 122.

<sup>37</sup> Gettier, "Is Justified," 123.

<sup>38</sup> Ralph H. Johnson and J. Anthony Blair, *Logical Self-Defense* (New York: International Debate Education Association, 2006), 154.

<sup>39</sup> Assuming that Jones used to own a Ford at one point in the past. If Jones has never owned a Ford, even in the past, then Smith's evidence would be misleading, and (iv) would again be false.

Blair, “equivocation occurs when the same word or phrase undergoes an unsignaled shift in meaning during one piece of discourse or argument.”<sup>40</sup> In Gettier’s Case II, the reference of ‘Jones’ in (i)–(iii) was fixed at some particular time in the *past*, since (i)–(iii) are based on what Smith *remembers* about Jones, whereas ‘Jones’ in (iv) is supposed to pick out the *present* Ford owner. This switch in reference “during one piece of discourse or argument” makes Gettier’s Case II appear like a genuine counterexample to JTB, even though it is not.

#### 4. Conclusion

In conclusion, other than engage in intuition mongering, Atkins does not really provide reasons to think that Gettier cases are not misleading. As I have argued in “Why Gettier Cases Are Misleading” and above, Gettier’s original cases are misleading because the candidates for knowledge in these cases contain ambiguous designators. In other words, in Gettier’s original cases, there is a mismatch between what the subjects wish to talk about (i.e., speaker’s reference) and what makes the relevant propositions true (i.e., semantic reference). In Atkins’ revised version of Gettier’s Case I, the ambiguous designator is ‘someone’. When Smith believes that there is someone who is both getting the job and is handsome, Smith has Jones in mind, for Smith’s evidence is about Jones, not about Smith himself. Indeed, Atkins himself admits that “there is some sense in which Smith has Jones in mind when inferring (I\*).”<sup>41</sup> In Gettier’s Case II, the ambiguity is in terms of the time to which the relevant beliefs are indexed. When Smith believes that either Jones own a Ford or Brown is in Barcelona, Smith has *past* Jones in mind, for Smith’s evidence is about *past* Jones, not about *present* Jones. So, again, there is a sense in which Smith has past Jones in mind when inferring (v) from (iv) by addition. For some reason that he does not specify, however, Atkins deems these ambiguities “irrelevant.”<sup>42</sup>

If this is correct, then it is still the case that Gettier cases are misleading because the candidates for knowledge in such cases contain ambiguous designators, which means that the relevant beliefs in such cases lend themselves to two interpretations: “an ‘objective’ interpretation in terms of the conditions that make the belief true (i.e., in terms of semantic reference or what a speaker’s words mean) and a ‘subjective’ interpretation in terms of what *S* means (i.e., in terms of speaker’s reference or what a speaker means in uttering certain words).”<sup>43</sup> Because

---

<sup>40</sup> Johnson and Blair, *Logical Self-Defense*, 154.

<sup>41</sup> Atkins, “Are Gettier Cases,” 383.

<sup>42</sup> Atkins, “Are Gettier Cases,” 383.

<sup>43</sup> Mizrahi, “Why Gettier Cases,” 43.

of this ambiguity, we should not assign much, if any, evidential weight to the so-called “Gettier intuition,” i.e., the intuition that  $S$  doesn’t know that  $p$  in a Gettier case.



# WEIGHING THE AIM OF BELIEF AGAIN

Asbjørn STEGLICH-PETERSEN

**ABSTRACT:** In his influential discussion of the aim of belief, David Owens argues that any talk of such an ‘aim’ is at best metaphorical. In order for the ‘aim’ of belief to be a genuine aim, it must be weighable against other aims in deliberation, but Owens claims that this is impossible. In previous work, I have pointed out that if we look at a broader range of deliberative contexts involving belief, it becomes clear that the putative aim of belief *is* capable of being weighed against other aims. Recently, however, Ema Sullivan-Bissett and Paul Noordhof have objected to this response on the grounds that it employs an undefended conception of the aim of belief not shared by Owens, and that it equivocates between importantly different contexts of doxastic deliberation. In this note, I argue that both of these objections fail.

**KEYWORDS:** belief, aim, epistemic norms, deliberation, propositional attitudes

## 1. Weighing the Aim of Belief

Many have been attracted to the idea that belief ‘aims’ at truth, in the hope of thereby demarcating belief from other propositional attitudes, and of explaining a number of puzzling features of belief, including the standard of correctness and epistemic norms governing belief. However, in his influential discussion of the aim of belief, David Owens argues that any talk of such an ‘aim’ is at best metaphorical.<sup>1</sup> In order for the ‘aim’ of belief to be a genuine aim, it must be weighable against other aims in deliberation. But Owens claims that this is impossible: when we deliberate over whether to believe some proposition, only truth-relevant considerations can have a say, to the exclusion of other kinds of considerations. No belief is ever the result of deliberative weighing of the aim of truth with other non-truth relevant aims and considerations. Belief does therefore not ‘aim’ at truth in a genuine and non-metaphorical sense that can carry its intended explanatory burden.

In my previous discussion of this argument, I pointed out that if we broaden our focus to other deliberative contexts involving belief, it becomes clear that the

---

<sup>1</sup> David Owens, “Does Belief Have an Aim?” *Philosophical Studies* 115 (2003): 283-305.

putative aim of belief *is* capable of the sort of weighing required of genuine aims.<sup>2</sup> In particular, when we deliberative over whether to *take up* the truth-aim with respect to some proposition, it is both relevant and possible to weigh it against other kinds of aims. For example, a teacher might weigh the aim of believing the truth as to which of her pupils broke the window against the aim of avoiding the unpleasant task of having to scold the guilty pupil. Such weighing may very well result in the teacher deciding *not* to pursue the aim of believing the truth with respect to that proposition. We might add that it is also possible for such considerations to enter into deliberation over whether to believe some particular proposition; no belief can result from such weighing, but it can cause the deliberation to be terminated without resulting in a belief. The reason that deliberation over whether to believe some proposition *p* does not allow weighing in a way that results in a belief as to whether *p*, I argued, is that such deliberation is essentially constrained by the aim of believing *p* if and only if *p* is true. This excludes the relevance of other kinds of considerations, except to convince one to give up the aim and terminate the deliberation. We can thus compare such deliberation to other similarly constrained examples of deliberation, such as deliberation over whether to go to some restaurant as a way of carrying out the aim of going there if and only if it received good reviews.

This explanation assumes, of course, that the aim one might take up as a result of deliberating whether to pursue the truth-aim with respect to some proposition *p* is the very aim that constrains deliberation over whether to believe that *p*, and is responsible for the resulting attitude being a belief. But I provided several examples to show that, on reflection, this assumption is quite plausible.<sup>3</sup> This does not entail that all beliefs are related to intentional aims in this way. As I have argued in another context, the aim of belief can be realized both by intentional aims of believers, and by sub-intentional mechanisms that share certain features with intentional aims.<sup>4</sup> Nor does it entail that the aim constraining deliberation over whether to believe that *p* is always the result of prior deliberation over whether to take up that aim. As with any other aim, it may or may not be the result of a deliberative process.

---

<sup>2</sup> Asbjørn Steglich-Petersen, "Weighing the Aim of Belief," *Philosophical Studies* 145 (2009): 395-405.

<sup>3</sup> Steglich-Petersen, "Weighing the Aim," 403-404.

<sup>4</sup> Asbjørn Steglich-Petersen, "No Norm Needed: On the Aim of Belief," *The Philosophical Quarterly* 56 (2006): 499-516.

## 2. Sullivan-Bissett and Noordhof's Reply

Ema Sullivan-Bissett and Paul Noordhof claim that this response to Owens' argument fails.<sup>5</sup> They advance two points in defense of Owens. Their first point<sup>6</sup> is that the examples I rely on fail because they invoke an undefended conception of the truth-aim not shared by Owens, and, they say, 'officially eschewed' by myself.<sup>7</sup> Owens construes the truth-aim as that of believing  $p$  only if  $p$  is true, thus making truth a necessary but not a sufficient condition for adopting belief. His reason for preferring this construal is to avoid implausibly attributing to believers the aim of believing each and every true proposition, however trivial. Sullivan-Bissett and Noordhof argue that if this is how we should understand the truth-aim, my examples of the truth-aim being weighed do not work, since in that case, the truth-aim does not insist on the agent forming any beliefs at all, and it therefore doesn't require any consideration whether or not to adopt this aim. For example, since adopting the truth-aim with respect to which of the pupils broke the window leaves the teacher free not to form any belief at all, it does not require any consideration or weighing against other aims on her behalf in deciding whether or not to adopt the aim.

There are several things to say in response to this argument. First of all, it is unclear why Sullivan-Bissett and Noordhof claim the if-and-only-if conception of the truth-aim to be 'officially eschewed' by myself. I am quite explicit in the discussion that I operate with this conception,<sup>8</sup> and it plays an integral and obvious role in my theory of doxastic deliberation, both in the article under discussion and elsewhere.<sup>9</sup> My guess is that they take the rejection of this conception as implied by me not objecting explicitly upon presenting Owens' conception, and his reason for preferring this. But nowhere else in the paper do I operate with Owens' conception.

Secondly, although I do not explicitly defend my own conception of the truth-aim in the paper, it should be clear that Owens' reason for preferring his conception is irrelevant on my account. Owens' reason was that we shouldn't

---

<sup>5</sup> Ema Sullivan-Bissett and Paul Noordhof, "A Defence of Owens' Exclusivity Objection to Beliefs Having Aims," *Philosophical Studies* 163 (2013): 453-457.

<sup>6</sup> Sullivan-Bissett and Noordhof, "A Defence of Owens," 455.

<sup>7</sup> Sullivan-Bissett and Noordhof, "A Defence of Owens," 453.

<sup>8</sup> E.g. on pages 402 and 404 in Steglich-Petersen, "Weighing the Aim."

<sup>9</sup> See e.g. Asbjørn Steglich-Petersen, "Voluntarism and Transparent Deliberation," *South African Journal of Philosophy* 25 (2006):171-176; Steglich-Petersen, "No Norm Needed"; and Asbjørn Steglich-Petersen, "Does Doxastic Transparency Support Evidentialism?" *Dialectica* 62 (2008): 541-547.

attribute to believers the aim of coming to a true belief with respect to all propositions. But this consideration only carries weight if we conceive of the truth-aim as a general aim in the first place. Clearly, believers don't have the aim of coming to a true belief *for any p*. But on my account, when believers have the aim of truth, they have it with respect to particular propositions or classes of propositions, not all propositions. So I do not attribute to believers a general aim of the sort rightly rejected by Owens.

Third, it clearly doesn't undermine my discussion that Owens doesn't share my preferred construal of the truth-aim. What is at issue is whether or not there is an interesting and non-metaphorical sense in which belief aims at truth, and in particular whether this aim satisfies Owens' requirement that it must be weighable against other aims. Owens (and Sullivan-Bissett and Noordhof) may have shown that on one particular construal of the truth-aim as a matter of necessary conditions for belief, this aim fails to satisfy Owens' requirement. But I can see no reason why it shouldn't be fair to object that there is *another* interesting construal of the truth aim that *does* satisfy Owens' weighing requirement.

Finally, it is all but clear that Sullivan-Bissett and Noordhof's argument holds, even if we accept that the truth-aim should be understood as a mere necessary condition for adopting belief (which I don't). It is certainly not in general the case that conditional aims of doing something *only if* some other condition obtains do not require and allow for weighing with other aims and considerations. Suppose, for example, that I am considering whether to aim for going to staff meetings *only if* there will be cake. Pursuing this aim could easily conflict with other aims of mine, such as the aim of staying on good terms with my Department Chair, and it is certainly relevant to weigh the cake-aim against this other aim in deliberation. It might be objected that such weighing is relevant only if one is interested in going to staff meetings in the first place: if one doesn't have any intention of going to staff meetings anyways, it would be a mute point whether one resolves only to go to meetings with cake. But that also seems too strong. Even if I am undecided on whether to go to staff meetings, it could still require weighing and consideration whether I should aim to go only if there will be cake. These considerations seem to apply to the belief case as well: even if the teacher does not yet have any intention of forming a belief as to which of the pupils broke the window, it could be a relevant matter for weighing and deliberation whether she should aim to adopt some belief on the matter only if the belief is true.



Sullivan-Bissett and Noordhof's second point<sup>10</sup> is that deliberating about whether to take up the truth-aim with respect to some proposition *p* is different from deliberating about whether to believe that *p*. As they say, Owens' claim about the exclusive relevance of truth was only meant to apply to the latter kind of deliberation, so I am missing the target when pointing out that the truth-aim is weighable in the former kind of deliberation. But I have never claimed that these two kinds of deliberation are the same; in fact, my main observation is that there are *several different* contexts of deliberation in which the truth-aim can play a role, and that it is weighable in at least one of these contexts. As I make explicit, my account assumes that the aim one might take up as a result of deliberating whether to pursue the truth-aim with respect to some *p*, is the *very aim* that constrains deliberation over whether to believe that *p*, thereby explaining the exclusive relevance of truth in this kind of deliberation.<sup>11</sup> This assumption is not beyond question, of course, but Sullivan-Bissett and Noordhof do not address it.

### 3. Conclusion

I conclude that Sullivan-Bissett and Noordhof's defense of Owens' exclusivity objection fails. Their first point rests on a misinterpretation of my conception of the truth-aim (and even if their interpretation had been correct, it is not clear that their point would survive). Their second point fails to address the idea that the aim one might take up as a result of deliberating whether to pursue a true belief as to whether *p*, can constrain deliberation over whether to believe that *p*.

---

<sup>10</sup> Sullivan-Bissett and Noordhof, "A Defence of Owens," 455-456.

<sup>11</sup> Steglich-Petersen, "Weighing the Aim," 403-404.



# ANOTHER DEFENCE OF OWENS'S EXCLUSIVITY OBJECTION TO BELIEFS HAVING AIMS

Ema SULLIVAN-BISSETT and Paul NOORDHOF

**ABSTRACT:** David Owens objected to the truth-aim account of belief on the grounds that the putative aim of belief does not meet a necessary condition on aims, namely, that aims can be *weighed* against other aims. If the putative aim of belief cannot be weighed, then belief does not have an aim after all. Asbjørn Steglich-Petersen responded to this objection by appealing to other deliberative contexts in which the aim could be weighed, and we argued that this response to Owens failed for two reasons. Steglich-Petersen has since responded to our defence of Owens's objection. Here we reply to Steglich-Petersen and conclude, once again, that Owens's challenge to the truth-aim approach remains to be answered.

**KEYWORDS:** belief, aim, truth, deliberation

## 1. Common Ground

Let us identify the common ground from which we and Steglich-Petersen begin. First, it is a necessary condition on aims that they are weighable.<sup>1</sup> Second, doxastic deliberation (deliberation over *whether to believe that p*) exhibits exclusivity to truth considerations<sup>2</sup> (indeed, Steglich-Petersen<sup>3</sup> has appealed to the aim of belief in explaining why this is so). The putative aim of belief then is not weighable in the context of doxastic deliberation. On these two points, all parties agree.

---

<sup>1</sup> To our knowledge, of all the responses to Owens's objection, no one has taken issue with this condition on aims.

<sup>2</sup> Conor McHugh has responded to Owens's objection by denying that deliberation over what to believe involves exclusivity to truth considerations, and thus the aim of belief can indeed be weighed, *in that very context* (Conor McHugh, "The Illusion of Exclusivity," *European Journal of Philosophy* 23 (2013): 1117-1136). We also think this response to Owens is unsuccessful (see Sophie Archer, "Exclusivity Defended," *Philosophy and Phenomenological Research* (2015), doi: 10.1111/phpr.12268; and Ema Sullivan-Bissett, "Aims and Exclusivity," *European Journal of Philosophy* (2017), DOI: 10.1111/ejop.12183).

<sup>3</sup> Asbjørn Steglich-Petersen, "Does Doxastic Transparency Support Evidentialism?" *Dialectica* 62, 4 (2008): 541-547, 546.

## 2. Steglich-Petersen's Reply (2009<sup>4</sup>) to Owens (2003<sup>5</sup>)

In reply, Steglich-Petersen identified other deliberative contexts in which the truth-aim can be weighed. He gave examples of the truth-aim being weighable insofar as it can be discarded in the context of deliberation over *whether to form a belief about  $p$*  (that is, whether to enter doxastic deliberation over *whether  $p$* ). Considerations speaking in favour of discarding the truth-aim and not forming a belief about *whether  $p$*  might be ones relating to the cognitive resources one is willing or able to devote to the task, or the consequences which might follow from forming a belief about  $p$ .

## 3. Our Response (2013<sup>6</sup>)

In response we made two points. First, in the cases Steglich-Petersen discusses, we should not say that the agents weigh the truth-aim and discard it in favour of other considerations, rather, the truth-aim does not require consideration. This is because agents are not required by the truth-aim to form beliefs, rather, it is only that *if* that is what an agent is up to, *then* the beliefs which she comes to have had better be true ones. This was captured by Owens's formulation of the truth-aim in terms of truth being a necessary (but not sufficient) condition for forming a belief that  $p$ .

Second, Steglich-Petersen equivocates between deliberating over *whether to form a belief about  $p$* , and deliberating over *whether to believe that  $p$* . In the former context, there is no adoption of the truth-aim. If an agent decides not to form a belief about  $p$ , Steglich-Petersen claims that she discards the truth-aim. But this is incorrect. Rather, the decision not to form a belief about  $p$  has been informed by a cognitive process prior to that of belief formation. The agent has not already adopted the truth-aim for  $p$ , and so is not weighing one aim against another. Truth is only a constraint upon what proposition one believes *if* one is in the business of forming a belief about a subject matter.

---

<sup>4</sup> Asbjørn Steglich-Petersen, "Weighing the Aim of Belief," *Philosophical Studies* 145 (2009): 395-405.

<sup>5</sup> David Owens, "Does Belief Have an Aim?" *Philosophical Studies* 115 (2003): 283-305.

<sup>6</sup> Ema Sullivan-Bissett and Paul Noordhof, "A Defence of Owens' Exclusivity Objection to Beliefs Having Aims," *Philosophical Studies* 163 (2013): 453-457.

#### 4. Steglich-Petersen's Counter (2017<sup>7</sup>) and Replies

Steglich-Petersen makes three points in reply to our previous defence of Owens's objection. Here, we take each in turn and offer a response, before making a final point.

##### 4.1. If and only if

Owens characterized the truth-aim as one in which the truth of *p* was *necessary* (but not *sufficient*) for belief that *p*. This was so as not to attribute to believers the aim of believing all true propositions. We noted that Steglich-Petersen accepts Owens's characterization of the truth-aim, and eschews the *if and only if* conception.<sup>8</sup>

In reply, he claims that it is unclear why we took the *if and only if* conception to be eschewed by him.<sup>9</sup> To this we note that if Steglich-Petersen was operating with the *if and only if* conception in his response to Owens, it was dialectically strange to characterize the truth-aim as Owens does, draw on it, and then fail to note that Owens had not characterized it in the right way, and that an alternative conception of the truth-aim would be operated with instead. Putting matters of interpretation aside though, Steglich-Petersen is now clear that he characterizes the truth-aim as one which has truth as *sufficient* for belief that *p*, but as not being vulnerable to the worry raised by Owens. This invulnerability is down to believers having the aim *with respect to particular propositions, or classes of propositions*, and not having a *general aim* of believing all true propositions.

Previously we prefigured a way of responding to us which had structurally similar features to Steglich-Petersen's position without talking of aims for local sets of propositions. We noted that any move to 'if and only if' might be meant not "as part of an alternative formulation of the truth aim but rather a description of what the agent is up to – given that it is now settled for him or her that he or she will arrive at a belief concerning whether or not the proposition is true."<sup>10</sup>

This conception of the truth-aim must pave a middle way between believers aiming to believe all the propositions which are true, and believers aiming to have only true beliefs. This middle way is restricted (so as to rule out the best avoided *truth as sufficient* construal), but is more liberal than applying only to those propositions subjected to doxastic deliberation (to rule out the *only if* construal).

---

<sup>7</sup> Asbjørn Steglich-Petersen "Weighing the Aim of Belief Again," *Logos & Episteme*, this issue.

<sup>8</sup> Sullivan-Bissett and Noordhof, "A Defence of Owens," 454.

<sup>9</sup> Steglich-Petersen "Weighing Again," 144.

<sup>10</sup> Sullivan-Bissett and Noordhof, "A Defence of Owens," 454.

The thought is that the aim kicks in for those propositions *one considers adopting the truth-aim for*, and then, the truth-aim can be weighed insofar as it can be discarded if the agent decides not to subject the proposition(s) to doxastic deliberation.

This conception of the truth-aim might help Steglich-Petersen's position only if it is the *same aim* which is present in doxastic deliberation, and the deliberative context which precedes it. (Later we suggest, contra Steglich-Petersen, that it is not, see §4.3, and even if it is, that does not yet do the work, see §4.4).

#### 4.2. Conditional Aims and Weighing

For the sake of argument, Steglich-Petersen grants that the truth-aim is best construed in terms of truth as merely *necessary* for belief. But he says, in cases of so-called *conditional aims* (where one aims to  $\phi$  only if some other condition obtains), it is not the case that other aims and considerations cannot be weighed against the conditional aim. He gives the example of the aim *to go to staff meetings only if there will be cake*, and suggests that pursuing this aim may conflict with one's other aims (e.g., maintaining good relations with the Department Chair), and these other aims are relevant in deliberation over whether to adopt the cake-aim.

This case is not to the point, since it was no part of our claim that conditional aims cannot be weighed. We were not taking issue with the *structure* of the truth-aim (its being conditional), but with the examples Steglich-Petersen used to demonstrate contexts in which that aim was *weighed*. We do not deny that there can be consideration of whether to adopt conditional aims; the truth-aim does indeed share with other conditional aims that a context preceding the aim's adoption can involve deliberation over whether to adopt the aim. Nevertheless, there is an important difference between the truth-aim and other conditional aims, which, we take it, is the basis of Owens's original concern. With conditional aims, it is possible to adopt the aim, the relevant condition for  $\phi$ -ing *not* be met, and yet  $\phi$  nevertheless. If I adopt the aim to *go to staff meetings only if there is cake*, it is possible for me to decide to go *even though there is not cake* (perhaps the meeting is especially important). Or if I adopt the aim to *run only if it is sunny*, it is possible for me to decide to run, *even if it is not sunny* (perhaps I am training for a marathon) (examples can be multiplied). But the analogous situation is ruled out in the case of the truth-aim: if I take up the truth-aim for some proposition  $p$  and enter into deliberation over *whether to believe that  $p$* , once I answer the question *whether  $p$*  in the negative or even fail to answer it in

the positive (and so the condition for belief is not met), I *cannot* form the belief *that p*. There is no parallel possibility of going ahead in spite of the condition not being met in the case of belief.

Steglich-Petersen is aware of this feature of belief formation but maintains it is still appropriate to characterize it as guided by the aim of truth. But given the difference identified, he needs to explain why this difference – not present in other cases – does not threaten the substantial use of aim talk here. We give reasons for supposing it does below (§4.4).

### 4.3. Equivocation

We argued that deliberating over *whether to form a belief about p* (to adopt the truth-aim for *p*) is not part of the belief-forming process. That if an agent is deliberating over *whether to be guided by the truth-aim*, she is not yet in the business of forming a belief. Deliberating over whether to adopt the aim with respect to a particular proposition is not a context in which one is already being guided by that aim.

Steglich-Petersen's claim is that the truth-aim is present in other deliberative contexts, and in some of those (such as whether to adopt that aim), the aim can be weighed. He notes that "the aim one might take up as a result of deliberating whether to pursue the truth aim with respect to some *p*, is the *very aim* that constrains deliberation over whether to believe that *p*."<sup>11</sup> So in the context of deliberating over *whether to form a belief about p*, one is working with the aim of belief (alongside others), in a stage prior to the belief-formation process.

We agree that deliberation on the question *whether to form a belief about p* gives way to practical considerations, and that *if* the truth-aim were present in such deliberation, *then* it would be weighed (perhaps against considerations regarding time or effort). But as we argued previously, deliberation over *whether to adopt the aim of  $\phi$ -ing* takes place in a context prior to  $\phi$ -ing, and so does not (perhaps *cannot*) involve the aim of  $\phi$ -ing itself. That is not to say one cannot adopt the aim of  $\phi$ -ing without actually beginning the process of  $\phi$  ing (aims can be dropped upon further reflection), but only to say that the deliberative context in which one considers *whether to adopt the aim to  $\phi$* , is not one in which *that very aim* plays a role.

Compare deliberating over *whether to adopt the aim of running only if it is sunny outside*. Deliberation over *whether to adopt this aim* will presumably include considerations of time, effort, injury-proneness, and so on. But to say that

---

<sup>11</sup> Steglich-Petersen "Weighing Again," 145.

this deliberative context is one in which the running-aim *itself* plays a role is, it seems to us, implausible, and at the very least, in need of argument.

Here is one way of thinking about the disagreement here. Steglich-Petersen's truth-aim can be read as a description of what the agent is up to once it is settled for her which proposition(s) to subject to doxastic deliberation (a point we made previously with respect to the possibility that Steglich-Petersen might opt for an *if and only if* aim).<sup>12</sup> Suppose that for a set of propositions  $S_1$ , the agent aims to *believe that p if and only if p is true*. Now we can ask, how is the set itself chosen – that is, how does the agent decide to adopt the truth aim for  $S_1$  instead of  $S_2$ ? Say that  $S_1$  is a set of propositions about what the weather will be like today, and  $S_2$  is a set of propositions about which student broke the classroom window. Presumably what settles one's adopting the 'if and only if' truth-aim towards one of these sets is a matter of Owens's formulation (if you are to form beliefs about the propositions in  $S_1$  they had better be true), and practical considerations. If it is important to my goals to form a true belief about what the weather will be like today, and/or if it would be troubling to form a true belief about who broke the classroom window, I might decide to believe all and only true propositions in  $S_1$ . But now we see that Steglich-Petersen's truth aim is not *weighed* against anything else, rather what we have is Owens's aim plus practical interests interacting. And this, of course, takes place at a stage prior to Steglich-Petersen's aim playing a role (if it does, see §4.4).

Our original charge was that Steglich-Petersen *equivocated* between deliberating over *whether to form a belief about p* and deliberating over *whether to believe that p*. He responded by noting that there is no such equivocation, that these are indeed different deliberative contexts, but that the truth-aim is at work in both of them (and weighed in the former). We replied here that the truth-aim is not at play in the context of deliberation over whether to adopt that exact aim and towards which proposition(s).

#### 4.4 Weighing Simpliciter Is Not the Point

As a final point, even if the truth aim were being weighed in a context prior to that of doxastic deliberation, that hardly shows that *belief formation* is governed by an aim. Rather, all that would be shown is that there is a truth-aim that can be weighed, and that (perhaps independently) belief formation follows rules that enable beliefs to be true. Just because a certain aim is weighed in deciding whether to deliberate over *p* does not mean that *that aim* is adopted in belief formation. It

---

<sup>12</sup> Sullivan-Bissett and Noordhof, "A Defence of Owens," 454.



could just be that the process of belief formation is such that its outputs are *in accordance with the aim*. When the process of deliberation is going on, the process no longer has the distinctive feature of being guided by the aim (after all, we might take guidance by the aim to be revealed by its being weighed against other aims). Consider the rules of a game. The rules may have been chosen to make the game pleasurable. But when one buys in and follows the rules, one does not allow pleasure to be weighed against other things. Likewise, on what grounds does Steglich-Petersen take bare weighability of the truth-aim in some context to support the claim that *that aim* is what structures the nature of some other context, that of belief formation?

## 5. Conclusions

We have again defended Owens's objection to beliefs having aims, this time from Steglich-Petersen's replies to our previous work on this issue. We agreed with Steglich-Petersen that the adoption of conditional aims can be preceded by deliberation over whether to adopt the aim in question, but argued that – unlike other conditional aims – one cannot go ahead and believe if the condition for believing specified by the aim is not met. One cannot ignore the prescription of the aim of belief. Indeed, it is this which motivated Owens's objection in the first place.<sup>13</sup>

We argued that if the truth aim is to be characterized as having the truth of *p* as sufficient for belief that *p*, *that aim* needs to be present in deliberation over whether to take up the truth aim, for there to be a case of the aim being weighed. But this is not so; deliberation over whether to adopt the aim is prior to deliberation structured by that aim, Steglich-Petersen has mistaken the interaction of the *only if* truth-aim with other interests for the weighing of the *if* and *only if* truth-aim.

Finally, we claimed that even if Steglich-Petersen were right that the truth aim *is* weighed in a context prior to belief formation, that does not show that belief formation itself is governed by an aim.

Thus we claim, again, that Owens's challenge to the truth-aim approach remains to be answered.

---

<sup>13</sup> For other explanations of why the prescription of the truth-aim cannot be ignored see Paul Noordhof, "Believe What You Want," *Proceedings of the Aristotelian Society* 101 (2001): 247–265; and Ema Sullivan-Bissett, "Explaining Doxastic Transparency: Aim, Norm, or Function?" *Synthese* (forthcoming). For limits to the prescription of the truth-aim see Paul Noordhof, "Self-Deception, Interpretation and Consciousness," *Philosophy and Phenomenological Research* LXVII, 1 (2003): 75–100.



# INFERENCES, EXPERIENCES, AND THE MYTH OF THE GIVEN: A REPLY TO CHAMPAGNE

Thomas WILK

**ABSTRACT:** In a recent article in this journal, Marc Champagne leveled an argument against what Wilfrid Sellars dubbed 'the Myth of the Given.' Champagne contends that what is given in observation in the form of a sensation must be able to both *cause* and *justify* propositionally structured beliefs. He argues for this claim by attempting to show that one cannot decide which of two equally valid chains of inference is sound without appeal to what is given in experience. In this note, I show that while this argument is sound, the conclusion he draws is far too strong. Champagne's argument shows only that our empirical beliefs are determined through experience. It does not license the stronger claim that, in order for us to have empirical knowledge, bare sensations must be able to justify beliefs.

**KEYWORDS:** the Myth of the Given, observation, justification, experience, empirical knowledge

## Introduction

In a recent article in this journal, Marc Champagne leveled an argument against what, in 1956, Wilfrid Sellars dubbed 'the Myth of the Given.'<sup>1</sup> In attacking Sellars's argument that the Given is a myth, Champagne also attacks a school of thought that follows in Sellars's footsteps most notably represented by Robert Brandom and John McDowell.<sup>2</sup> Champagne contends that what is given in observation in the form of a sensation can, indeed must, both *cause* and *justify* propositionally structured beliefs. He argues for this claim by attempting to show that one cannot decide which of two equally valid chains of inference is sound without appeal to what is given in experience. In this note, I show that while this

---

<sup>1</sup> Marc Champagne, "Tracking Inferences Is Not Enough: The Given as Tie-Breaker," *Logos & Episteme* 7, 2 (2016): 129-135; Wilfrid Sellars, "Empiricism and the Philosophy of Mind," in *Knowledge, Mind, and the Given: Reading Wilfrid Sellars's "Empiricism and the Philosophy of Mind,"* eds. Willem A. Devries and Timm Triplett (Indianapolis: Hackett Publishing Company, 2000), 205-276.

<sup>2</sup> Chauncey Maher, *The Pittsburgh School of Philosophy: Sellars, McDowell, Brandom* (London: Routledge, 2012).

argument is sound, the conclusion he draws is far too strong. Champagne's argument shows only that our empirical beliefs are determined through experience, but this is something that no one denies – Sellars, Brandom, and McDowell included. His argument does not license the stronger claim that, in order for us to have empirical knowledge, bare sensations must be able to justify beliefs.

### What is 'the Given'?

Let's start by setting the bar for the success of Champagne's argument: what would it have to show in order to refute the Sellarsian claim that the Given is a myth? To answer this, we need to know what Sellars meant in labelling the Given as such. His primary concern is with a foundationalist picture of knowledge insofar as it takes all knowledge – both of particulars and of general empirical truths – to rest on a stratum of cognitive states that are both epistemically independent and epistemically efficacious.<sup>3</sup> This picture requires that these cognitive states – sensations, sensings, knowledge of sense data, seeings – be epistemically efficacious for the obvious reason that if they are not, then they cannot pass on whatever positive epistemic status they have to any further cognitive states. These basic cognitive states must be able to support the edifice of empirical knowledge. The picture requires that they be epistemically independent – that they have their positive epistemic status independent of their relationship to other cognitive states – because, if they were not, they could not serve as true foundations. If they presuppose knowledge of other particular matters of fact or general empirical truths then they cannot, by themselves, serve as the tribunal against which further empirical claims are tested. Cognitive states that are both epistemically independent and epistemically efficacious are 'the Given.'

Now, Sellars's claim is that no cognitive state can have both of these characteristics. This is what makes the Given a myth, and a pernicious one at that. In order for any cognitive state to be epistemically efficacious, it must be propositionally structured. This follows from the nature of inference: only propositionally structured contents can stand in inferential relations to one another. Knowledge of sense data is ruled out on these grounds.

Sellars argues further that cognitive states with propositionally structured content are not epistemically independent. This argument proceeds by cases, but one example should be sufficient to get its flavor. A classic proposal for filling in

---

<sup>3</sup> Sellars, "Empiricism," sec. VIII; also see Willem deVries, "Wilfrid Sellars," in *The Stanford Encyclopedia of Philosophy* (2016), ed. Edward N. Zalta, sec. 4, <<http://plato.stanford.edu/archives/fall2016/entries/sellars/>>.

the strata of basic empirical beliefs in the foundationalist picture appeals to sentences similar in form to “There looks to be a physical object with a red and triangular facing surface.”<sup>4</sup> *Lookings* or *appearings* seem to be just what are needed to provide a firm footing for empirical knowledge since (1) the concepts invoked in an *appearing* have a plausible claim to epistemic independence and (2) though one can be wrong about what one sees, one cannot be mistaken about how things appear to her. This incorrigibility is appealing, but Sellars argues that it is the product not of an ability to report on some *minimal, objective* facts but of withholding full endorsement of the propositional content of the claim. “[T]he statement ‘X looks green to Jones’ differs from ‘Jones sees that x is green’ in that whereas the latter both ascribes a propositional claim to Jones’s experience *and endorses it*, the former ascribes the claim but does not endorse it.”<sup>5</sup> This undermines (2), but if Sellars is right, then this also entails that the notion of *being green* is not reducible to that of *looking green*, for “the ability to recognize that x looks green presupposes the concept of *being green*.”<sup>6</sup> This means that (1) is also called into question since the ability to use the concept of *being green* presupposes knowledge of what circumstances count as standard conditions for observing colors and an ability to determine whether those circumstances obtain, which presupposes knowledge of a range of other perceptibles besides. Looks talk, though epistemically efficacious, is not epistemically independent, and so cannot serve as ‘the Given.’

Champagne’s argument would have to do one of three things in order to convince us that the Given is not a myth. (1) He might propose by way of example some item that is given in experience that is both epistemically efficacious and epistemically independent. This would involve the construction of an entire epistemology of perception, but his article is not nearly so ambitious. (2) He might show that Sellars’s arguments are somehow badly mistaken. Champagne does ask his readers to recall “that philosophers who reject the given do so, not in response to some tangible crisis, but on account of a technical let-down: it is not propositional, and therefore cannot enter into an argument.”<sup>7</sup> This, however, is not the main thrust of his argument, and he does not develop the thought in any detail. Finally, (3) he might show that it is necessary for something to be given in

---

<sup>4</sup> Sellars, “Empiricism,” sec. 9.

<sup>5</sup> *Ibid.*, sec. 16; Robert B. Brandom, *From Empiricism to Expressivism: Brandom Reads Sellars* (Cambridge: Harvard University Press, 2015), 105-109; Devries and Triplett, ed., *Knowledge, Mind, and the Given*, chap. 3.

<sup>6</sup> Sellars, “Empiricism,” sec. 19.

<sup>7</sup> Champagne, “Tracking Inferences,” 133-134.

experience in order for us to have any empirical knowledge at all. Givenness might be the cost of avoiding skepticism. This would be odd since the Cartesian desire for firm foundations is at the root of external-world skepticism, but this seems to be the course Champagne pursues. He aims to argue that encounter with a bare given in experience is necessary for empirical knowledge. Let's turn now to this argument.

### Champagne's Müller-Lyer Illusion Argument

We are asked to consider the Müller-Lyer illusion, a simple visual illusion in which arrowheads are appended to the ends of two parallel lines of equal length. On one line, the arrowheads point inward, on the other outward. To the observer, this setup gives the illusion that the line with outward facing arrowheads is longer than that with inward facing arrowheads. Champagne asks us to imagine a naïve observer sitting in a darkened room. She is unfamiliar with the illusion. It is described to her in sufficient detail, and then she is given the following argument:

1) The Müller-Lyer lines appear uneven

2) The Müller-Lyer lines are even

3) Illusions are not as they appear

Therefore,

4) The Müller-Lyer lines are an illusion<sup>8</sup>

Does the subject *know* the conclusion of this argument? We are to assume that she knows what all the terms mean and that she grasps the inferential relations being laid before her. We could also assume that this naïve observer grasps many of the other inferences adjacent to this particular sequence. She might grasp, for example, that undertaking a commitment to the claim that the two lines are uneven would commit her to the further claim that if one were to draw perpendicular lines at the ends of the 'longer' line, these newly drawn lines would pass by the ends of the shorter line without touching it and that undertaking a commitment to the claim that the two lines are even would entail that, in performing the same operation, the perpendicular lines would make contact at both ends of both lines.<sup>9</sup> The ability to do this – to draw out inferences entailed by a commitment one undertakes – is enough to credit this observer with

---

<sup>8</sup> *Ibid.*, 131.

<sup>9</sup> *Ibid.*, 132.

rationality, but none of this entails that she *knows* the conclusion of the argument.<sup>10</sup>

Champagne's point is that this inferential chain remains idle unless one is given 'an observational cause' to affirm the first premise. "Reasoning alone might establish the formal validity of the inference presented in the darkness, but the only way for the subject to assess the soundness of the argument is for her to take advantage of the experiential deliverances which alone can establish whether the first premise is true."<sup>11</sup>

Now, there is a way in which this is already too quick. The soundness of the inference could be secured by testimony. If our observer has reason to trust the account of the illusion given to her, then she might accept the conclusion on the authority of the explainer. Let's set aside testimony, however, for we must admit that though much of our knowledge rests on testimony, the edifice of empirical knowledge cannot on the whole. At some point, observation must play a role, and this is the point that Champagne is keen to make.

So, observation is necessary to, as Champagne puts it, break the tie between two equally valid chains of inference: a *modus ponens* establishing the truth of the conclusion and one *establishing* its falsity. Does this show that the Given isn't a myth after all since it is required for empirical cognition? No. This would follow only if Sellars, Brandom, and McDowell understood the myth of the Given as an argument against the possibility of any perceptual encounters with the world licensing claims to knowledge. Sellars's argument in "Empiricism and the Philosophy of Mind" doesn't do this. Rather, it shows that there is a particular shape that such encounters cannot take: they cannot be cognitive states that are both epistemically independent and epistemically efficacious. To put this another way, it shows that bare encounters with the world cannot provide a rational constraint on our thinking *without* calling into play certain capacities that belong to our conceptual apparatus.<sup>12</sup> This is far from saying that perceptual encounters with the world are impossible or unimportant for empirical knowledge.

Champagne's argument shows us only that the deliverances of perceptual experience are required in order to break the tie between the two potential chains of inference, but he has nothing at all to say about what shape perceptual experience must take. He claims only that whether the Müller-Lyer lines are or appear even must be "ascertained by looking" and that "claims and inferences are

---

<sup>10</sup> *Ibid.*, 133.

<sup>11</sup> *Ibid.*, 132.

<sup>12</sup> John McDowell, *Mind and World* (Cambridge: Harvard University Press, 1994), 66.

answerable to the experiential qualities before one.”<sup>13</sup> Neither Sellars, Brandom, nor McDowell reject the need to ground empirical knowledge in experiential encounters with the world; all three, and other Sellarsians besides, agree that we require a theory of non-inferential knowledge. They are in accord with Champagne on this point. The problem, though, is to develop such a theory without falling afoul of the myth. While much of the work of these three authors is devoted to just this problem, Champagne ignores it entirely.

### Sellarsians on Perception

Sellars, for his part, develops a positive epistemology that is part reliabilist and part internalist.<sup>14</sup> He argues first that we must possess dispositions to reliably respond differentially to perceptual stimuli. These reliable differential responsive dispositions (RDRDs, for Brandom) are something genuine knowers like us share with all sentient critters. What separates us from them in terms of epistemic abilities is that we have the capacity to reliably differentially respond *by applying concepts*. Our responses are perceptual judgments. Applying concepts (and, so, making judgments), for Sellars, is a matter of mastering the use of words, which involves the ability to take up a position in the game of giving and asking for reasons. In particular, applying concepts in perceptual judgment involves undertaking a commitment to the content of that judgment as something that can both stand in need of and serve as a reason. It is making oneself liable to give reasons for the judgment and committing oneself to its downstream consequences. This is where the internalist component comes in, for a reliable responsive disposition to differentially apply concepts can count as a judgment only if one knows that one’s RDRDs are indeed reliable. It is only if this is the case that one could give reasons for the perceptual judgment to which one has undertaken a commitment.<sup>15</sup>

Brandom develops this Sellarsian position in a social pragmatic direction. There is a problem lurking in Sellars’s account: one can have perceptual knowledge only if one knows that one’s RDRDs are reliable, but it seems that one could only come to know *that* on the basis of past experiences of their reliability. The problem is that those experiences couldn’t have counted as instances of

---

<sup>13</sup> Champagne, “Tracking Inferences,” 133.

<sup>14</sup> Sellars, “Empiricism,” sec. 35.

<sup>15</sup> deVries, “Wilfrid Sellars,” sec. 4; Robert Brandom, *Tales of the Mighty Dead: Historical Essays in the Metaphysics of Intentionality* (Cambridge: Harvard University Press, 2002), chap. 12; Paul Coates, *The Metaphysics of Perception: Wilfrid Sellars, Perceptual Consciousness and Critical Realism* (Oxford: Taylor & Francis, 2009).



perceptual knowledge.<sup>16</sup> Brandom's solution is to argue that the agent herself need not make the reliability inference. It is the knowledge attributor who attributes reliability to the knower. This is a route adjacent to strong internalism. It recognizes that *someone* must recognize that the reporter is, in fact, reliable, but takes that burden off of the reporter herself. Knowledge is not just accidentally, but necessarily, a social phenomenon.<sup>17</sup>

Finally, there is McDowell. In *Mind and World*, he characterizes the myth of the Given as an episode in the 'interminable oscillation' between a picture of perception that has no place for receptivity and one in which the recognized need for external constraint on empirical thought motivates us to reintroduce the Given, i.e., between coherentism and foundationalism.<sup>18</sup> Both poles of this oscillation are problematic. On the one hand, conceptual thought – the product of pure spontaneity – fails to be constrained by contact with the world. We are left with a picture of "the operations of spontaneity as a frictionless spinning in the void."<sup>19</sup> On the other hand, when we take the Given to provide the needed external constraint, we have a picture of pure receptivity in which the conceptual capacities of spontaneity are wholly absent. This gives us only the illusion of external constraint, for, as Sellars argued, non-conceptual cognitive states cannot be epistemically efficacious. McDowell's response to this oscillation is to argue for a middle ground, an understanding of experience as at once passive and drawing "into operation the capacities that genuinely belong to spontaneity."<sup>20</sup>

---

<sup>16</sup> Rebecca Kukla develops an intriguing account of these past experiences being "constitutively misremembered" in order to solve this problem. Rebecca Kukla, "Myth, Memory and Misrecognition in Sellars' 'Empiricism and the Philosophy of Mind,'" *Philosophical Studies: An International Journal for Philosophy in the Analytic Tradition* 101, 2-3 (2000): 161-211.

<sup>17</sup> Robert Brandom, *Making It Explicit: Reasoning, Representing, and Discursive Commitment* (Cambridge: Harvard University Press, 1994), chap. 4; Robert Brandom, "Knowledge and the Social Articulation of the Space of Reasons," *Philosophy and Phenomenological Research* 55, 4 (1995): 895-908; also see Rebecca Kukla and Mark Lance, *"Yo!" and "Lo!": The Pragmatic Topography of the Space of Reasons* (Boston: Harvard University Press, 2009), chap. 2. Kukla and Lance develop an account of the pragmatics of perception and argue that the Sellarsian tradition errs in thinking that observational episodes must be propositionally structured. They claim that this follows from too narrow a construal of inference and that, in the end, what is required is that they be conceptual.

<sup>18</sup> McDowell, *Mind and World*, 9.

<sup>19</sup> *Ibid.*, 11.

<sup>20</sup> *Ibid.*, 13.

### **Conclusion**

McDowell's diagnosis of the oscillation furnishes a lens through which to view Champagne's argument. Champagne understands the rejection of the Given as a myth as synonymous with the endorsement of the opposite pole of oscillation, namely, coherentism. His worry is that in rejecting the Given we confine thought to a frictionless spinning in the void or, perhaps worse, thought becomes paralyzed. When faced with equally valid inferences issuing in contradictory conclusions, we have no reason for endorsing one over the other without some encounter with the world through experience. Champagne's response is to recoil to the other pole. This is precisely the mistake Sellars hoped to warn us against. As our excursion into Sellarsian territory has shown, each of Champagne's targets recognizes this demand for external constraint. The myth of the Given is not the rejection of experience as a source of knowledge. Recognition of the myth requires, however, that we accept certain constraints on how we understand experience. It cannot be a bare, non-conceptual encounter with the world if it is to be epistemically efficacious.

Champagne closes by claiming, "Givenness, whatever else it might be, is the tie-breaker," but as I believe I have shown, it is perceptual experience, not the Given, that breaks the tie. The problem with which we are faced is how to conceive of such experiences without falling afoul of the myth. In the end, Champagne is right that tracking inferences is not enough, but neither Sellars nor later Sellarsians thought that it was.

# NOTES ON THE CONTRIBUTORS

**Fred Adams** is Professor of Linguistics & Cognitive Science and Professor of Philosophy at the University of Delaware. He has published well over 100 articles and chapters on cognitive science, epistemology, philosophy of language, philosophy of action, and related topics. He is also co-author of *The Bounds of Cognition* (Wiley-Blackwell, 2008, with Ken Aizawa) and of *An Introduction to the Philosophy of Psychology* (Cambridge University Press, 2015, with Daniel Weiskopf. Contact: fa@udel.edu.

**Guy Axtell** is a Professor of Philosophy, and a Faculty Fellow of the Radford University Honors Academy. His main teaching interests are in epistemology and metaphysics, philosophy of science, “STS” or science, technology and society studies, and philosophy of religion. His most recent publications include: “William James on Pragmatism and Religion,” and “The Emotions in James’ *Principles of Psychology*” (in *William James, Moral Philosophy, and the Ethical Life: The Cries of the Wounded*, ed. Jacob Goodson, Lexington Books, 2016), *Objectivity* (Polity Press, 2015), and “Navigating the Dialectics of Objectivity” (in *The Future of Social Epistemology: A Collective Vision*, ed. James Collier, Rowman & Littlefield, 2015). Contact: gsaxtell@radford.edu.

**John A. Barker** is Professor Emeritus of Philosophy at Southern Illinois University Edwardsville. He is the author of numerous articles in the areas of epistemology, philosophy of language, logic, philosophy of mind, and philosophy of science. He is also the author of a book, *A Formal Analysis of Conditionals* (Southern Illinois University Monographs, 1969) and of a software application, *ProtoThinker: A Model of the Mind* (Wadsworth, 1998 and ProtoThinker.com, 2016). Contact: johnabarker1@cs.com.

**Umut Baysan** is a postdoctoral research associate in Philosophy at the University of Glasgow. His works mainly in metaphysics and philosophy of mind, and has research interests in epistemology and philosophy of science. His publications include “Realization Relations in Metaphysics” (*Minds and Machines*, 2015), “An Argument for Power Inheritance” (*Philosophical Quarterly*, 2016), and book reviews in *Mind* and *Philosophical Quarterly*. He is also the author of the

“Realization” entry in *Oxford Bibliographies in Philosophy*. Contact: emin.baysan@glasgow.ac.uk.

**John Biro** teaches at the University of Florida. His interests are in the philosophy of language, the philosophy of mind, epistemology, metaphysics and the history of philosophy, especially in the seventeenth and eighteenth centuries. He has published papers in these areas in journals such as *The Monist*, *Australasian Journal of Philosophy*, *Philosophy and Phenomenological Research*, *Philosophical Studies*, *History of Philosophy Quarterly*, *The Review of Metaphysics* and *Hume Studies*. He is also co-editor of two collections of papers on Spinoza and one on Frege. Contact: jbiro@ufl.edu.

**Olga Ramirez Calle** is Assistant Professor of Philosophy at the University of Saint Louis Missouri, Campus Madrid. She obtained her PhD (Dr. Phil) at the J.W. Goethe University of Frankfurt at Main. Her research interests are principally concerned with foundational questions and lie at the crossroads between philosophy of language, epistemology and metaphysics. She has worked on Crispin Wright’s conception of truth as superassertibility, metaethics (specifically conceptual genealogy) and has defended the existence of ‘functionally structured concepts’ and the expansion of Russell’s quantified existential analysis to functions to give account of them. She has also written a study on the grounds of logical inference, which is ongoing work. Contact: olgaramirezcallegmail.com.

**Murray Clarke** is Professor in the Department of Philosophy at Concordia University in Montreal, Quebec. His research interests span contemporary epistemology, cognitive science and philosophy of science. He is the author of *Reconstructing Reason and Representation* (MIT, 2004). He has published numerous articles including “Resurrecting the Tracking Theories” (*Australasian Journal of Philosophy*, 2005), “Concepts, Intuitions and Epistemic Norms” (*Logos & Episteme*, 2010), and “Critical Notice: Jose Zalabardo’s Scepticism and Reliable Belief” (*Teorema*, 2014). Contact: Murray.Clarke@concordia.ca.

**Jonathan L. Kvanvig** is Professor of Philosophy at Washington University in St. Louis. With extensive publications in areas such as epistemology, philosophy of religion, logic, and philosophy of language, he is also the administrator of the epistemology blog *Certain Doubts*, the editor of *Oxford Studies in Philosophy of Religion*, and the organizer of the Annual Philosophy of Religion Conference. His most recent books are *Faith and Humility* (Oxford University Press, 2018) and

*Rationality and Reflection* (Oxford University Press, 2014). Contact: jkvanvig@wustl.edu.

**Kevin McCain** is Assistant Professor of Philosophy at the University of Alabama at Birmingham. His primary areas of research are epistemology and philosophy of science. He is the author of two books: *Evidentialism and Epistemic Justification* (Routledge, 2014) and *The Nature of Scientific Knowledge* (Springer, 2016). He is currently co-editing, with Ted Poston, *Best Explanations: New Essays on Inference to the Best Explanation* (Oxford University Press, forthcoming) and *The Mystery of Skepticism: New Explorations* (Brill, forthcoming). Contact: mccain@uab.edu.

**Moti Mizrahi** is an Assistant Professor of Philosophy at the Florida Institute of Technology. He has teaching and research interests in argumentation theory, epistemology, ethics, metaphysics, philosophy of religion, and philosophy of science. He has published or forthcoming papers in journals such as *Argumentation*, *Axiomathes*, *European Journal for Philosophy of Religion*, *Informal Logic*, *International Journal for the Study of Skepticism*, *International Studies in Philosophy of Science*, *Journal for General Philosophy of Science*, *Metaphilosophy*, *Philosophia*, *Ratio*, *Review of Philosophy and Psychology*, *Social Epistemology*, *Teorema*, and *Think*. Contact: motimizra@gmail.com.

**Paul Noordhof** is Anniversary Professor of Philosophy at the University of York. His main research interests are in the philosophy of mind, action theory, and metaphysics. His work in the philosophy of mind mainly focuses upon the nature and explanatory character of consciousness. His publications include “Believe What You Want” (*Proceedings of the Aristotelian Society*, 2001), “Self-Deception, Interpretation and Consciousness” (*Philosophy and Phenomenological Research*, 2003); “The Essential Instability of Self-Deception” (*Social Theory and Practice*, 2009); and “A Defence of Owens’ Exclusivity Objection to Beliefs Having Aims” (*Philosophical Studies*, 2013, with Ema Sullivan-Bissett). Contact: paul.noordhof@york.ac.uk.

**Asbjørn Steglich-Petersen** is Associate Professor in Philosophy at Aarhus University. He has published widely in epistemology, philosophy of mind, metaphysics, and philosophy of language, with a major strand of work devoted to epistemic normativity and the nature of belief. His articles have appeared in journals such as *Mind*, *The Philosophical Quarterly*, *Synthese*, *Erkenntnis*, *Mind*

*and Language, Philosophical Studies, Journal of Semantics, British Journal of the Philosophy of Science, Utilitas, and in edited volumes published by Oxford and Cambridge University Press. He is also co-editor of the volume *Reasons for Belief* (Cambridge University Press, 2011). Contact: [filasp@cas.au.dk](mailto:filasp@cas.au.dk).*

**Emma Sullivan-Bissett** is a Lecturer in Philosophy at the University of Birmingham. Her research interests are in the philosophy of mind and psychology, specifically belief and its connection to truth, as well as self-deceptive beliefs and delusional beliefs. She is also interested in biological approaches to what are characteristically thought to be normative questions in philosophy of mind, epistemology, and ethics. Her publications include “A Defence of Owens’ Exclusivity Objection to Beliefs Having Aims’ (*Philosophical Studies*, 2013, with Paul Noordhof), “Aims and Exclusivity” (*European Journal of Philosophy*, forthcoming), “Biological Function and Epistemic Normativity” (*Philosophical Explorations*, forthcoming); and “Malfunction Defended” (*Synthese*, 2016). Contact: [e.l.sullivan-bissett@bham.ac.uk](mailto:e.l.sullivan-bissett@bham.ac.uk).

**Thomas Wilk** is a PhD candidate in Philosophy at Johns Hopkins University and an adjunct instructor at various colleges and universities around Washington, DC. His research interests are primarily in metaethics, philosophy of language, epistemology, and applied ethics, but he also works on the philosophy of human rights, the history of early modern philosophy, and issues in the study of race, gender, and social justice. Contact: [wilkth01@gmail.com](mailto:wilkth01@gmail.com).

## ***LOGOS & EPISTEME: AIMS & SCOPE***

*Logos & Episteme* is a quarterly open-access international journal of epistemology that appears at the end of March, June, September, and December. Its fundamental mission is to support philosophical research on human knowledge in all its aspects, forms, types, dimensions or practices.

For this purpose, the journal publishes articles, reviews or discussion notes focused as well on problems concerning the general theory of knowledge, as on problems specific to the philosophy, methodology and ethics of science, philosophical logic, metaphilosophy, moral epistemology, epistemology of art, epistemology of religion, social or political epistemology, epistemology of communication. Studies in the history of science and of the philosophy of knowledge, or studies in the sociology of knowledge, cognitive psychology, and cognitive science are also welcome.

The journal promotes all methods, perspectives and traditions in the philosophical analysis of knowledge, from the normative to the naturalistic and experimental, and from the Anglo-American to the Continental or Eastern.

The journal accepts for publication texts in English, French and German, which satisfy the norms of clarity and rigour in exposition and argumentation.

*Logos & Episteme* is published and financed by the "Gheorghe Zane" Institute for Economic and Social Research of The Romanian Academy, Iasi Branch. The publication is free of any fees or charges.

For further information, please see the *Notes to Contributors*.

Contact: [logosandepisteme@yahoo.com](mailto:logosandepisteme@yahoo.com).





# NOTES TO CONTRIBUTORS

## 1. Accepted Submissions

The journal accepts for publication articles, discussion notes and book reviews.

Please submit your manuscripts electronically at: [logosandepisteme@yahoo.com](mailto:logosandepisteme@yahoo.com). Authors will receive an e-mail confirming the submission. All subsequent correspondence with the authors will be carried via e-mail. When a paper is co-written, only one author should be identified as the corresponding author.

There are no submission fees or page charges for our journal.

## 2. Publication Ethics

The journal accepts for publication papers submitted exclusively to *Logos & Episteme* and not published, in whole or substantial part, elsewhere. The submitted papers should be the author's own work. All (and only) persons who have a reasonable claim to authorship must be named as co-authors.

The papers suspected of plagiarism, self-plagiarism, redundant publications, unwarranted ('honorary') authorship, unwarranted citations, omitting relevant citations, citing sources that were not read, participation in citation groups (and/or other forms of scholarly misconduct) or the papers containing racist and sexist (or any other kind of offensive, abusive, defamatory, obscene or fraudulent) opinions will be rejected. The authors will be informed about the reasons of the rejection. The editors of *Logos & Episteme* reserve the right to take any other legitimate sanctions against the authors proven of scholarly misconduct (such as refusing all future submissions belonging to these authors).

## 3. Paper Size

The articles should normally not exceed 12000 words in length, including footnotes and references. Articles exceeding 12000 words will be accepted only occasionally and upon a reasonable justification from their authors. The discussion notes must be no longer than 3000 words and the book reviews must not exceed 4000 words, including footnotes and references. The editors reserve the right to ask the authors to shorten their texts when necessary.

#### **4. Manuscript Format**

Manuscripts should be formatted in Rich Text Format file (\*.rtf) or Microsoft Word document (\*.docx) and must be double-spaced, including quotes and footnotes, in 12 point Times New Roman font. Where manuscripts contain special symbols, characters and diagrams, the authors are advised to also submit their paper in PDF format. Each page must be numbered and footnotes should be numbered consecutively in the main body of the text and appear at footer of page. For all references authors must use the Humanities style, as it is presented in The Chicago Manual of Style, 15th edition. Large quotations should be set off clearly, by indenting the left margin of the manuscript or by using a smaller font size. Double quotation marks should be used for direct quotations and single quotation marks should be used for quotations within quotations and for words or phrases used in a special sense.

#### **5. Official Languages**

The official languages of the journal are: English, French and German. Authors who submit papers not written in their native language are advised to have the article checked for style and grammar by a native speaker. Articles which are not linguistically acceptable may be rejected.

#### **6. Abstract**

All submitted articles must have a short abstract not exceeding 200 words in English and 3 to 6 keywords. The abstract must not contain any undefined abbreviations or unspecified references. Authors are asked to compile their manuscripts in the following order: title; abstract; keywords; main text; appendices (as appropriate); references.

#### **7. Author's CV**

A short CV including the author's affiliation and professional postal and email address must be sent in a separate file. All special acknowledgements on behalf of the authors must not appear in the submitted text and should be sent in the separate file. When the manuscript is accepted for publication in the journal, the special acknowledgement will be included in a footnote on the first page of the paper.

## **8. Review Process**

The reason for these requests is that all articles which pass the editorial review, with the exception of articles from the invited contributors, will be subject to a strict double anonymous-review process. Therefore the authors should avoid in their manuscripts any mention to their previous work or use an impersonal or neutral form when referring to it.

The submissions will be sent to at least two reviewers recognized as specialists in their topics. The editors will take the necessary measures to assure that no conflict of interest is involved in the review process.

The review process is intended to take no more than six months. Authors not receiving any answer during the mentioned period are kindly asked to get in contact with the editors. Processing of papers in languages other than English may take longer.

The authors will be notified by the editors via e-mail about the acceptance or rejection of their papers.

The editors reserve their right to ask the authors to revise their papers and the right to require reformatting of accepted manuscripts if they do not meet the norms of the journal.

## **9. Acceptance of the Papers**

The editorial committee has the final decision on the acceptance of the papers. Papers accepted will be published, as far as possible, in the order in which they are received and they will appear in the journal in the alphabetical order of their authors.

## **10. Responsibilities**

Authors bear full responsibility for the contents of their own contributions. The opinions expressed in the texts published do not necessarily express the views of the editors. It is the responsibility of the author to obtain written permission for quotations from unpublished material, or for all quotations that exceed the limits provided in the copyright regulations.

## **11. Checking Proofs**

Authors should retain a copy of their paper against which to check proofs. The final proofs will be sent to the corresponding author in PDF format. The author must send an answer within 3 days. Only minor corrections are accepted and should be sent in a separate file as an e-mail attachment.

## **12. Reviews**

Authors who wish to have their books reviewed in the journal should send them at the following address: Institutul de Cercetări Economice și Sociale „Gh. Zane” Academia Română, Filiala Iași, Str. Teodor Codrescu, Nr. 2, 700481, Iași, România. The authors of the books are asked to give a valid e-mail address where they will be notified concerning the publishing of a review of their book in our journal. The editors do not guarantee that all the books sent will be reviewed in the journal. The books sent for reviews will not be returned.

## **13. Property & Royalties**

Articles accepted for publication will become the property of *Logos & Episteme* and may not be reprinted or translated without the previous notification to the editors. No manuscripts will be returned to their authors. The journal does not pay royalties.

## **14. Permissions**

Authors have the right to use their papers in whole and in part for non-commercial purposes. They do not need to ask permission to re-publish their papers but they are kindly asked to inform the Editorial Board of their intention and to provide acknowledgement of the original publication in *Logos & Episteme*, including the title of the article, the journal name, volume, issue number, page number and year of publication. All articles are free for anybody to read and download. They can also be distributed, copied and transmitted on the web, but only for non-commercial purposes, and provided that the journal copyright is acknowledged.

## **15. Electronic Archives**

The journal is archived on the Romanian Academy, Iasi Branch web page. The electronic archives of *Logos & Episteme* are also freely available on Philosophy Documentation Center web page.